Nederlandse Zorgautoriteit

Tussenrapport

Kwantitatief onderzoek naar risicoselectie

Eerste inzichten

juni 2015

Inhoud

Vooraf

De Nederlandse Zorgautoriteit (NZa) geeft in haar toezicht op de zorgverzekeraars prioriteit aan het monitoren van de risicosolidariteit. Uit eerdere onderzoeken blijkt dat er geen aanwijzingen zijn dat zorgverzekeraars de acceptatieplicht of het verbod op premiedifferentiatie schenden, maar er wel signalen zijn dat er mogelijk sprake is van andere vormen van risicoselectie op de zorgverzekeringsmarkt. Ook als een vorm van risicoselectie niet in strijd is met de wet, druist het wel tegen de belangen in die de wet beoogt te beschermen, zoals de risicosolidariteit.

Bepaalde vormen van risicoselectie kunnen ook andere nadelige gevolgen hebben voor consumenten en voor de werking van het stelsel. Risicoselectie kan bijvoorbeeld leiden tot het verschraling van de zorg of innovatie in de zorg ontmoedigen. Een ander effect zou kunnen zijn dat zorgverzekeraars zich minder inspannen doelmatig zorg in te kopen.

Gezien het belang van risicosolidariteit en het betaalbaar houden van de zorg, is het essentieel dat de NZa de ontwikkelingen op de zorgverzekeringsmarkt op het punt van risicoselectie monitort zodat bij ongewenste ontwikkelingen maatregelen kunnen worden genomen. Dit tussenrapport is de eerste stap in een groter onderzoek dat risicoselectie monitort en toezichtacties en andere maatregelen kan ondersteunen.

De Nederlandse Zorgautoriteit,

dr. M.J. Kaljouw
voorzitter Raad van Bestuur

Managementsamenvatting

De NZa geeft in haar toezicht prioriteit aan het onderwerp risicoselectie. Naast onderzoeken naar onder andere de acceptatieplicht, sturing van patiëntenstromen en de relatie tussen de basis- en aanvullende zorgverzekering, wil de NZa ook een totaal beeld geven van mogelijke risicoselectie in de zorgverzekeringsmarkt. Risicoselectie is niet in strijd met de wet, maar druist wel in tegen de belangen die de wet beoogt te beschermen, zoals de risicosolidariteit. Risicosolidariteit houdt in dat mensen met een kleine kans op ziekte evenveel betalen als mensen die een slecht gezondheidsrisico hebben.
Bepaalde vormen van risicoselectie kunnen er ook toe leiden dat innovaties in de zorg ontmoedigd worden of doelmatigheid uit het oog wordt verloren. Om de ontwikkelingen op dit gebied goed te kunnen volgen en om mogelijk te maken dat bij ongewenste ontwikkelingen maatregelen kunnen worden genomen, wil de NZa risicoselectie op een structurele wijze monitoren. Inzichten over de aard en mate van risicoselectie kan helpen de risico's op nadelige gevolgen te verminderen of weg te nemen bijvoorbeeld door de prikkels voor risicoselectie te wijzigen (bijvoorbeeld via doorontwikkeling van het risicovereveningsmodel) of door regelgeving aan te scherpen.

Dit tussenrapport is de eerste stap in een groter onderzoek naar risicoselectie. Voor een kwantitatief onderzoek waarmee op een gestructureerde manier naar signalen wordt gezocht, zijn gedetailleerde gegevens nodig. Over de benodigde gegevens beschikte de NZa eerder niet. Het opvragen van deze gegevens en het uitbouwen van de geschikte infrastructuur die de koppeling van gegevens van verschillende bronnen mogelijk maakt, neemt veel tijd in beslag. Om tussenresultaten zo snel mogelijk te kunnen publiceren, hebben we besloten het onderzoek in meerdere fases uit te voeren. De fases zijn bepaald door de beschikbaarheid van de gegevens uit verschillende bronnen.

Dit rapport is een tussenrapportage waarin we de eerste (deel)resultaten presenteren. De analyses zijn uitgevoerd op basis van gegevens van het Zorginstituut Nederland. Zij maken de rapportage van resultaten mogelijk op het niveau van zorgverzekeraars (risicodragers). De gegevens die analyses op het niveau van polissen mogelijk maken, zijn niet bij één centrale bron beschikbaar maar moesten bij de afzonderlijke zorgverzekeraars opgevraagd worden. Omdat het verzamelen van deze aanvullende gegevens langer duurde, worden analyses op het niveau van polissen in de tweede fase van het onderzoek meegenomen. Het eindrapport van het onderzoek waarin deze aanvullende gegevens zijn verwerkt presenteren we naar verwachting in het vierde kwartaal van 2015.

In dit tussenrapport presenteren we een set basis indicatoren die we bij het monitoren van signalen van risicoselectie willen gebruiken. De beschreven methodologie hebben we vervolgens toegepast, zodat we een eerste inzicht hebben in de ontwikkelingen op de zorgverzekeringsmarkt. We gebruiken gegevens die ook in het risicovereveningssysteem worden gebruikt. De basis van onze indicatoren is het berekende vereveningsresultaat per verzekerde. Om de resultaten te kunnen interpreteren, is het noodzakelijk om een paar concepten uit het risicoverevening te definiëren.

Risicovereveningssysteem
De verplichting om iedereen te accepteren voor de basisverzekering is cruciaal voor het solidaire stelsel van de basisverzekering in Nederland. Zorgverzekeraars worden gecompenseerd via het zogenoemde risicovereveningssysteem voor voorspelbare kostenverschillen tussen verzekerden. Alle zorgverzekeraars krijgen uit het Zorgverzekeringsfonds jaarlijks een financiële bijdrage, die afhankelijk is van de persoonskenmerken (zoals leeftijd, geslacht, sociale economische status) van hun verzekerden. Bij elk persoonskenmerk dat gebruikt wordt in de verevening hoort een normbedrag. Het normbedrag geeft aan wat het marginale effect is van het persoonskenmerk op de totale kosten. Het verschil tussen het bedrag berekend in het risicovereveningssysteem en de werkelijke zorgkosten is het vereveningsresultaat.

Een negatief vereveningsresultaat betekent dan dat een zorgverzekeraar in een bepaald jaar hogere gemiddelde kosten per verzekerde heeft dan door het vereveningsmodel berekend wordt. Een positief resultaat houdt in dat een zorgverzekeraar lagere gemiddelde kosten per verzekerde heeft dan op basis van het vereveningsmodel berekend wordt. Wij hebben in dit onderzoek gekeken naar het vereveningsresultaat voor bepaalde subgroepen van verzekerden en het vereveningsresultaat per zorgverzekeraar.

Definitie van risicoselectie
In dit onderzoek gebruiken we de definitie van risicoselectie beschreven in het artikel 'Hoe kan de Nederlandse Zorgautoriteit (NZa) risicoselectie op de zorgverzekeringsmarkt aantonen en meten?'[1]

"acties (anders dan premiedifferentiatie per product) door verzekerden of verzekeraars met als doel of gevolg dat de beoogde risicosolidariteit niet volledig wordt gerealiseerd"

Deze definitie is breder dan de bewuste actie van een zorgverzekeraar om verzekerden met een gunstig risicoprofiel aan te trekken of verzekerden met een ongunstig risicoprofiel af te houden. Onder deze definitie vallen ook acties van zorgverzekeraars (bijvoorbeeld het aanbieden van een beperkte polis voor prijsgevoelige verzekerden) die een ander doel hebben, maar die als gevolg hebben dat de risicosolidariteit wordt aangetast. Ook kan zelfselectie door verzekerden, ertoe leiden dat de risicosolidariteit in gevaar komt. Deze twee effecten zijn moeilijk van elkaar te scheiden. In dit onderzoek gebruiken we de brede definitie van risicoselectie. In de tweede fase van het onderzoek proberen we de resultaten die we vinden te verklaren en de verschillende strategieën van zorgverzekeraars die tot risicoselectie leiden te onderscheiden.

Resultaten in dit tussenrapport
In dit tussenrapport hebben we alle analyses op het niveau van zorgverzekeraars (risicodragers) gedaan. Het vereveningsresultaat dat we berekenen betreft alleen de deelbijdrage somatische zorg van het risicovereveningssysteem. Doordat we de analyses over een langere periode doen, periode 2008-2012, krijgen we inzicht in de robuustheid van de bevindingen. De relevante vraag bij de robuustheid van de resultaten is of zorgverzekeraars met een positief vereveningsresultaat in één jaar erin slagen om dit resultaat structureel te behalen. Een structureel positief vereveningsresultaat betekent dat de verzekerden van de zorgverzekeraar elk jaar gemiddeld lagere kosten hebben dan

[1] Van de Ven, Van Kleef, Van Vliet (2015), in Tijdschrift voor Gezondheidswetenschappen, jaargang 93/ 2015 nummer 4.

door het risicovereveningsmodel voorspeld.[2] Omdat de kosten structureel lager zijn, is de kans klein dat het om een eenmalige, uitzonderlijke situatie gaat.

We zijn in onze berekeningen uitgegaan van de werkelijke kosten. Dat betekent dat we informatie gebruiken die pas drie jaar na het betreffende jaar beschikbaar komt. Het gevolg van het rekenen met werkelijke kosten is dat de effecten van onvoorspelbare kostenontwikkelingen en van aannames in het risicovereveningsmodel die achteraf niet juist blijken in onze resultaten meegenomen zijn. In een aparte indicator kwantificeren we wel het effect van de onvoorspelbare ontwikkelingen op het resultaat van zorgverzekeraars. Hieruit blijkt dat hoewel onvoorspelbare ontwikkelingen een effect hebben op het resultaat van zorgverzekeraars, zij niet bepalend zijn.

Op drie gebieden biedt dit tussenrapport aanknopingspunten voor verder onderzoek.

Ten eerste, het tussenrapport biedt inzichten over verschillen tussen subgroepen van verzekerden. In deze eerste analyse hebben we het vereveningsresultaat (voor somatische zorg) van drie subgroepen bekeken. De subgroep 'jongeren' had een licht negatief vereveningsresultaat in 2009 en 2010. In andere jaren lag het resultaat rond nul. Dit betekent dat jongeren voor de zorgverzekeraars in de somatische zorg in 2009 en 2010 gemiddeld meer kostten dan door het risicovereveningsmodel voorspeld was. De voorlopige conclusie is dat het niet gunstig was voor zorgverzekeraars om veel jongeren onder hun verzekerden te hebben in deze periode. Deze voorlopige conclusie kan nog veranderen door het meenemen van het resultaat in het ggz en eigen risico model in de tweede fase van het onderzoek.
Ouderen (65 jaar en ouder) daarentegen hadden in 2009 een hoog positief vereveningsresultaat op somatische zorg. Dat nam in de periode 2010-2012 geleidelijk af. Ongezonde ouderen[3] hadden zelfs een fors positief gemiddeld vereveningsresultaat. Een positief vereveningsresultaat betekent dat de werkelijke gemiddelde kosten van deze groep verzekerden lager uitkwamen in de somatische zorg dan het door het risicovereveningsmodel voorspeld was. De voorlopige conclusie is dat zorgverzekeraars die in deze periode veel ouderen en vooral ongezonde ouderen onder hun verzekerden hadden beter af waren dan zorgverzekeraars bij wie deze groepen ondervertegenwoordigd waren. Deze voorlopige conclusie kan nog veranderen in de tweede fase van het onderzoek door het meenemen van het resultaat in het ggz- en eigen risico model.

Ten tweede laat het tussenrapport verschillen tussen de vereveningsresultaten van afzonderlijke zorgverzekeraars (risicodragers) zien. We hebben het vereveningsresultaat van elke zorgverzekeraar berekend voor zijn totale populatie, de instroom en de uitstroom. Het beeld dat naar voren komt is dat het resultaat van ongeveer de helft van de zorgverzekeraars in een beperkte range schommelt. Deze zorgverzekeraars hadden in sommige jaren hogere en in andere jaren lagere gemiddelde kosten (in de somatische zorg) per verzekerde dan door het vereveningsmodel voorspeld werd. Hun resultaten lijken dus door eenmalige gebeurtenissen te worden bepaald en vormen in eerste instantie geen signaal voor risicoselectie. Het aantal uitschieters waarbij

---

[2] Omdat er steeds nieuwe kenmerken worden toegevoegd aan het vereveningsmodel om de voorspellende waarde van het model te verbeteren, verwachten we dat het vereveningsresultaat van elke zorgverzekeraar steeds dichterbij nul uitkomt.
[3] Ongezonde ouderen zijn verzekerden ouder dan 65 jaar die ingedeeld zijn een DKG, FKG of MHK risicoklasse (voor de definitie van deze risicoklassen zie hoofdstuk 3 Gegevens).

op basis van het resultaat op de totale populatie in eerste instantie sprake kan zijn van signalen voor risicoselectie is relatief klein. Dit kan het gevolg zijn van het aggregatieniveau van de analyse. Omdat de meeste zorgverzekeraars meerdere polissen hebben, middelen de resultaten van de afzonderlijke polissen van een zorgverzekeraar uit. Analyses op het niveau van polissen in de tweede fase van het onderzoek lossen dit probleem op en geven een beter beeld over signalen voor mogelijke risicoselectie.
Verder is het opvallend dat er meer variatie is in het resultaat op de instroom per zorgverzekeraar, dan in het resultaat op de totale populatie of de uitstroom. Dit lijkt erop te duiden dat er sprake is van risicoselectie doordat er een meer selecte groep verzekerden bij zorgverzekeraars instroomt dan de groep die ze hebben of de groep verzekerden die hen verlaat. Verder onderzoek op polisniveau (in de tweede fase van het onderzoek) is nodig om in kaart te brengen welke strategieën van een zorgverzekeraar een rol spelen in de selecte instroom van verzekerden.

Ten derde hebben we de normbedragen opnieuw geschat op basis van de werkelijke kosten van verzekerden. Hiermee krijgen we normbedragen die ook de effecten bevatten van onverwachte kostenontwikkelingen. De werkelijke normbedragen houden daar geen rekening mee. Het verschil tussen de opnieuw geschatte normbedragen en de werkelijke normbedragen geeft aan welke risicoklassen werden geraakt door onverwachte kostenontwikkelingen. Zoals we zien, zijn de risicoklassen met weinig verzekerden de uitschieters. Een meer gedetailleerde analyse van verschillen kan nuttige informatie bieden voor de bouwers van het risicovereveningsmodel.

Vervolg
De komende tijd breiden we dit onderzoek uit met analyses op het polisniveau en naar de vereveningsmodellen ggz en het eigen risico. In het eindrapport, dat naar verwachting in het vierde kwartaal van dit jaar verschijnt, onderzoeken we het statistische verband tussen het vereveningsresultaat van afzonderlijke polissen en de kenmerken van deze polissen zoals de polisvoorwaarden, hoogte van de korting voor het vrijwillig eigen risico en collectiviteitskortingen. Met statistische analyses willen we het verschil in het vereveningsresultaat tussen polissen verklaren, zodat we de mechanismen op de markt beter grijpen en risicoselectie kunnen voorkomen.

Het is echter niet mogelijk om alles, zoals de beweegredenen van zorgverzekeraars, te duiden op basis van kwantitatieve analyses. In dit onderzoek kunnen we ook niet de huidige situatie op de markt analyseren omdat de benodigde gegevens pas later beschikbaar komen. Daarom begint de NZa ook met een kwalitatief onderzoek waarin individuele casussen, die we op basis van signalen in dit onderzoek kiezen, worden toegelicht. Afhankelijk van de uitkomsten van het kwantitatief en kwalitatief onderzoek zal de NZa aanbevelingen doen en zo nodig en mogelijk maatregelen nemen.

1. Inleiding

De NZa houdt ook toezicht op de zorgverzekeraars en monitort het goed functioneren van de zorgverzekeringsmarkt. Het gegeven dat consumenten bewust een polis kunnen kiezen en vrij zijn om over te stappen naar een andere zorgverzekeraar is een belangrijk element in de marktwerking. Tegenhanger daarvan is dat zorgverzekeraars proberen verzekerden voor hun polis te winnen. Risicoverevening zorgt ervoor dat verzekerden met hoge of lage verwachte zorgkosten naar verwachting even duur zijn voor de zorgverzekeraars.[4] Perfecte risicoverevening bestaat echter niet. Er zijn altijd subgroepen van verzekerden te vinden voor wiens kosten de risicoverevening niet goed compenseert en die dus voorspelbaar winstgevend of juist verliesgevend zijn. In het geval van imperfecte risicoverevening bestaat er een prikkel voor de zorgverzekeraar om winstgevende verzekerden aan te trekken en verliesgevende verzekerden te weren. De NZa pleit er dan ook voor dit systeem actueel te houden en steeds aan te passen, zodat zorgverzekeraars niet kunnen voorspellen welke groepen verzekerden hun op langere termijn geld kosten of opleveren.

Zorgverzekeraars hebben een acceptatieplicht, dat wil zeggen dat iedereen ongeacht zijn achtergrond of ziektehistorie tegen dezelfde voorwaarden een basisverzekering moet kunnen afsluiten. Daardoor is risicoselectie door het weren van verzekerden niet mogelijk.[5] Er zijn wel andere vormen van risicoselectie die in beginsel niet in strijd zijn met de wet. Een zorgverzekeraar kan bijvoorbeeld zijn polis alleen onder die verzekerden adverteren die naar verwachting winstgevend zijn. Of hij kan een aanvullende verzekering op de markt zetten die precies deze groep aanspreekt. Het is ook mogelijk dat een zorgverzekeraar minder oog heeft in zijn inkoopbeleid voor de belangen van naar verwachting verliesgevende groepen waardoor zijn polissen niet aantrekkelijk worden voor deze (naar verwachting verliesgevende) groepen. Door risicoselectie komt de solidariteit in de zorgverzekering in het geding. Hoewel andere vormen van risicoselectie dan het weren van verzekerden formeel gezien niet in strijd zijn met de wet, druisen zij wel in tegen de belangen die de wet beoogt te beschermen. Wanneer zorgverzekeraars doelbewust acties uitvoeren om bepaalde groepen verzekerden aan te trekken en andere groepen af te houden, concurreren de zorgverzekeraars alleen op de verzekerdengroep die voor hen lucratief zijn en behartigen zij niet de belangen van verliesgevende verzekerdengroepen. Dit gedrag kan nadelige effecten hebben op de langere termijn.[6]

Het onderwerp risicoselectie wordt steeds actueler, omdat zorgverzekeraars steeds meer financiële risico's gaan dragen nu de ex-post compensaties geleidelijk worden afgebouwd.[7] Doordat zorgverzekeraars meer financieel risico lopen hebben zij meer prikkels om hun resultaten te verbeteren met behulp van risicoselectie en/ of doelmatige zorginkoop. Terwijl doelmatige zorginkoop een gewenste uitkomst van de marktwerking is, is risicoselectie dit niet. Het is dus

[4] Voor verdere uitleg over het risicovereveningssysteem zie hoofdstuk 2.

[5] Daarnaast geldt er een verbod op premiedifferentiatie naar aan de persoon van de verzekerde gerelateerde factoren. Het verbod op premiedifferentiatie draagt bij aan het borgen van risicosolidariteit. Premiedifferentiatie is echter per definitie geen vorm van risicoselectie.

[6] Voor een opsomming van alle gevolgen van risicoselectie zie het rapport 'Hoe kan de NZa risicoselectie meten?' http://www.nza.nl/104107/138040/iBMG_-_Onderzoeksrapport_Hoe_kan_de_NZa_risicoselectie_meten__iBMG-april_2013.pdf.

[7] Een uitleg over het risicovereveningssysteem staat in hoofdstuk 2.

essentieel om de zorgverzekeringsmarkt te monitoren op signalen van risicoselectie en indien nodig in te grijpen.

De NZa geeft in haar toezicht prioriteit aan het onderwerp risicoselectie. In onze marktscans over de Zorgverzekeringsmarkt 2013 en 2014 besteedden we daar aandacht aan. In 2013 is het instituut Beleid & Management Gezondheidszorg (iBMG, Erasmus Universiteit Rotterdam) gevraagd om te onderzoeken hoe de NZa kan monitoren of er sprake is van risicoselectie en om de voorgestelde kwantitatieve indicatoren ook toe te passen[8]. De NZa heeft de uitkomsten van deze onderzoeken gerapporteerd in de Marktscan Zorgverzekeringsmarkt 2013. In 2014 heeft de NZa een onderzoek uitgevoerd naar de naleving van de acceptatieplicht en sturing van verzekerdenstromen door zorgverzekeraars.[9] De NZa heeft in 2014 ook samengewerkt met de Autoriteit Consument en Markt (ACM) om consumenten te ondersteunen in het maken van een keuze voor een zorgverzekering voor 2015. Verder heeft de NZa samen met de ACM en Autoriteit Financiële Markten (AFM) onderzoek gedaan naar de werkwijze en informatievoorziening van vergelijkingssites. In vervolg hierop deed de NZa in het voorjaar van 2015 verdiepend onderzoek naar de naleving van de acceptatieplicht door volmachten en naar de vraag wanneer er sprake is van ongewenste vormen van sturing van verzekerden voor de basisverzekering via de aanvullende verzekering. We presenteren de bevindingen van beide onderzoeken tegelijkertijd met dit tussenrapport.

Naast deze onderzoeken wil de NZa een totaal beeld krijgen van risicoselectie in de zorgverzekeringsmarkt. Om de ontwikkelingen op dit gebied goed te kunnen volgen en om mogelijk maken dat bij ongewenste ontwikkelingen maatregelen kunnen worden genomen wil de NZa mogelijke risicoselectie op een structurele wijze monitoren. In dit tussenrapport presenteren we een set van indicatoren en de eerste inzichten die zij opleveren. Later dit jaar werken we de hier beschreven analyses verder uit in een uitgebreidere eindrapport. Door dit onderzoek de komende jaren periodiek te herhalen kunnen we de set van indicatoren steeds verder verfijnen.

1.1 Definitie

Er is sprake van risicoselectie als zorgverzekeraars door middel van een bewuste actie, zoals bepaalde marketingstrategieën, verzekerden met een gunstig risicoprofiel proberen te selecteren. Het is moeilijk om de effecten van verschillende strategieën, zoals doelmatige zorginkoop en risicoselectie, van elkaar te onderscheiden, en onmogelijk om de precieze beweegredenen van zorgverzekeraars te achterhalen. Het is dus niet met volledige zekerheid aan te tonen of de uitkomsten die we vinden met kwantitatieve analyses het resultaat zijn van bewuste acties van verzekeraars om hun verzekerden op risico te selecteren. Deze gebruikelijke interpretatie van risicoselectie valt daarom niet te operationaliseren.

[8] Zie het rapport in voetnoot 1 en het rapport 'Overstapgedrag en risicoselectie op de zorgverzekeringsmarkt' http://www.nza.nl/104107/138040/iBMG_Overstapgedrag_en_risicoselectie_op_de_zorgverzekeringsmarkt_augustus2013.pdf.

[9] Verdiepend onderzoek naleving acceptatieplichtplicht door zorgverzekeraars, april 2014, NZa.

In het artikel 'Hoe kan de Nederlandse Zorgautoriteit (NZa) risicoselectie op de zorgverzekeringsmarkt aantonen en meten' (Van de Ven et al, 2015)[10] geven de auteurs een bredere definitie van risicoselectie:

"acties (anders dan premiedifferentiatie per product) door verzekerden of verzekeraars met als doel of gevolg dat de beoogde risicosolidariteit niet volledig wordt gerealiseerd"

Deze definitie maakt geen onderscheid in de acties van de zorgverzekeraars en de verzekerden en kijkt ook niet naar het feit of het resultaat (geen volledige risicosolidariteit) het doel of het gevolg van de acties was. Onderstaande tabel geeft een typering van acties die onder deze definitie van risicoselectie vallen.

Tabel 1.1. Typering van acties

	Acties met als doel	Acties met als gevolg
Actie door verzekeraars	Type I Voorbeeld: niet inspelen op de preferenties van groepen chronisch zieken die worden ondergecompenseerd door de risicoverevening met als doel "deze groep buiten de deur te houden"	Type III Voorbeeld: verbeteren van de kwaliteit van zorg voor chronisch zieken met als neveneffect dat de verzekeraar extra chronisch zieken aantrekt en dus verlies maakt omdat deze chronische zieken worden ondergecompenseerd door de risicoverevening.
Actie door verzekerden	Type II Voorbeeld: gezonden kiezen een budgetpolis met beperkte keuzevrijheid met als doel de premie te minimaliseren.	Type IV Voorbeeld: keuze voor de beste zorg door chronisch zieken en keuze voor goede zorg door niet-chronisch zieken, waardoor deze groepen in verschillende polissen terechtkomen.

Bron: Van de Ven et al. (2015)

Acties type I komen overeen met de gebruikelijke interpretatie van risicoselectie. Zoals uit de voorbeelden blijkt kunnen verschillende type acties tegelijk voorkomen en samen voor een bepaalde uitkomst zorgen. In de praktijk betekent dit dat we niet kunnen vaststellen welke acties tot een bepaald resultaat leden.

Acties van het type II, III en IV leiden vooral tot de afname van risicosolidariteit, maar ze hebben verder minder negatieve effecten.[11] Acties van type I hebben meer negatieve effecten; ze kunnen naast de afname van solidariteit ook leiden tot afname van de kwaliteit en doelmatigheid van de zorg. (Van de Ven et al, 2015, p.7). Daarom zou het wenselijk zijn om type I acties te kunnen onderscheiden van de andere typen. Met kwantitatief onderzoek is dat echter niet mogelijk. Verdiepende (kwalitatieve) analyses kunnen meer helderheid bieden, maar het is niet mogelijk om met volledige zekerheid doelbewuste risicoselectie vast te stellen.

---

[10] Van de Ven, Van Kleef, Van Vliet (2015), in Tijdschrift voor Gezondheidswetenschappen, jaargang 93/ 2015 nummer 4.

[11] In het voorbeeld van het tabel lijkt actie type III zelfs een positief effect te hebben voor chronisch zieken (een polis op de markt die specifiek op de behoeften van chronisch zieken inspeelt). We verwachten echter dat deze situatie niet voorkomt in de markt omdat de zorgverzekeraar verlies zou lijden zoals dat ook in het voorbeeld staat.

Uit kwantitatieve gegevens zijn alleen de netto effecten van verschillende acties af te leiden. Dit betekent dat als zorgverzekeraars acties uitvoeren om risicoselectie onder verzekerden toe te passen die geen effect hebben, we die niet kunnen aantonen met kwantitatieve analyses. De uitkomsten van een kwantitatief onderzoek biedt daarom altijd een onderschatting van de aanwezigheid van risicoselectie.

In dit onderzoek gebruiken we de definitie van Van de Ven et al (2015) omdat de verschillende type acties niet met zekerheid te onderscheiden zijn. Door het gebruik van deze definitie sluiten we ons ook aan bij de internationale literatuur.[12] Bij de interpretatie van onze resultaten moet wel rekening worden gehouden met deze bredere definitie (zie ook paragraaf 1.3 'Interpretatie van resultaten') en met het feit dat als we geen risicoselectie signaleren dat niet noodzakelijk de afwezigheid van risicoselectie betekent. Het is essentieel om naast dit kwantitatieve onderzoek de ontwikkelingen op de polismarkt ook met kwalitatieve methodes te volgen.

1.2 Onderzoeksvragen

Het is onmogelijk om aan te tonen dat er geen sprake is van risicoselectie. Het is ook niet makkelijk om vast te stellen dat er wèl sprake is van risicoselectie, omdat risicoselectie niet los te zien is van andere strategieën van de zorgverzekeraars. Om de meest gevoelige casussen diepgaand te kunnen onderzoeken, noemt iBMG een aantal kwantitatieve toetsen in het rapport 'Hoe kan de NZa risicoselectie meten'. Met deze kwantitatieve toetsen kunnen we de markt monitoren en signalen genereren voor een verdiepend (kwalitatief) onderzoek. Deze kwantitatieve toetsen nemen we over in dit tussenrapport.

In dit tussenrapport onderzoeken we de volgende vragen:
- Hadden bepaalde subgroepen van de verzekerden (op landelijk niveau) een positief of negatief vereveningsresultaat[13] voor de somatische zorg in de periode 2008-2012?
- In welke mate hebben de afzonderlijke zorgverzekeraars een selectieve populatie of in- of uitstroom gehad in de periode 2008-2012 die leidde tot een positief of negatief vereveningsresultaat voor de somatische zorg?

Signalen voor risicoselectie worden op basis van de tweede onderzoeksvraag gegenereerd. De eerste onderzoeksvraag helpt in het begrijpen van mechanismen op de zorgverzekeringsmarkt. Het doel van dit onderzoek is niet alleen het signaleren van mogelijke risicoselectie maar ook handvatten te bieden voor het voorkomen van risicoselectie. Door het beter begrijpen van de mechanismen op de markt, kunnen we gerichter en effectiever maatregelen nemen bij ongewenste ontwikkelingen.

Bij de eerste vraag analyseren we subgroepen zowel op basis van kenmerken die al in het vereveningsmodel zijn opgenomen als op basis van combinaties van deze kenmerken. Als blijkt dat bepaalde kenmerken

[12] Bijvoorbeeld 'actions of economic agents on either side of the market to exploit unpriced risk heterogeneity and break pooling arrangements' (Newhouse, J.P., 1996, Reimbursing Health Plans and Health Providers: Selection versus Efficiency in Production , Journal of Economic Literature, 34, p.1236).

[13] Een negatief vereveningsresultaat betekent dan dat een zorgverzekeraar in een bepaald jaar hogere gemiddelde kosten per verzekerde had dan op basis van het verevening berekend werd. Een positief resultaat houdt in dat een zorgverzekeraar lagere gemiddelde kosten per verzekerde had dan op basis van het verevening berekend werd. (Voor verder uitleg over het risicovereveningssysteem zie hoofdstuk 2.)

waarvoor het vigerende model corrigeerde (ook na het toepassen van ex-post correcties) tot een hoog positief resultaat leidden, kan dit betekenen dat het model de uitkomsten op deze kenmerken niet goed kon voorspellen, bijvoorbeeld door onverwachte kostenontwikkelingen specifiek voor deze groep.[14] Het vereveningsresultaat was dan niet te voorspellen op basis van historische gegevens die in de risicoverevening worden gebruikt. Mocht een zorgverzekeraar onevenredig veel verzekerden met dit kenmerk hebben, dan vinden we een hoog vereveningsresultaat. Als zorgverzekeraars de relatief gunstige kosten van deze verzekerden niet tijdens de jaarlijkse overstapperiode konden voorspellen, zijn deze echter niet het gevolg van risicoselectie.

Een andere mogelijkheid is dat het model niet goed compenseert als bepaalde kenmerken samenvallen (bijvoorbeeld bij gezonde ouderen). Dit soort imperfecties in het model geven de mogelijkheid verzekerden met een gunstig risicoprofiel te selecteren. Het is dus belangrijk om te kijken welke groepen een structureel positief of negatief vereveningsresultaat hebben en deze aan de uitkomsten per zorgverzekeraar te relateren.

Bij de tweede onderzoeksvraag bekijken we indicatoren op het niveau van zorgverzekeraars (risicodragers) met de focus op het totale werkelijke resultaat; de instroom en uitstroom. De bevindingen worden aangevuld met een indicator die de effecten voortvloeiend uit bijvoorbeeld onvoorspelbare kostenontwikkelingen kwantificeert.

De indicatoren over de totale verzekerdenpopulatie en de in- en uitstroom volgen uit de aanbevelingen van iBMG. De indicatoren worden uitgewerkt en beschreven in het hoofdstuk 4 ‘Indicatoren’.

1.3 Interpretatie van de resultaten

Bij het interpreteren van resultaten is het belangrijk te benadrukken dat we doelen op het daadwerkelijke vereveningsresultaat dat achteraf bepaald wordt.[15] We gebruiken dus meer informatie dan waarover zorgverzekeraars beschikten op het moment waarover de analyse gaat. Bijvoorbeeld, in de overstapperiode 2010-2011 waren kostengegevens van verzekerden gedeeltelijk beschikbaar tot en met 2010. Daarvan waren de jaren 2009-2010 onvolledig. Zorgverzekeraars konden hun strategieën alleen op deze gegevens baseren. In onze analyses gebruiken we volledige gegevens over de werkelijke kosten van 2011. De resultaten zijn dus niet alleen het gevolg van acties van zorgverzekeraars (en verzekerden), maar ook van externe factoren waar niemand invloed op heeft. Het effect van deze niet voorspelbare kostenontwikkelingen kwantificeren we in een aparte indicator.

Het is ook de vraag in hoeverre zorgverzekeraars het verwachte vereveningsresultaat van verzekerden kunnen berekenen. Zorgverzekeraars kennen niet alle kenmerken van hun verzekerden die in het vereveningsmodel worden gebruikt. Zo weten ze de sociaal economische status en de aard van inkomen van hun verzekerden niet.

[14] Een voorbeeld hiervoor is de forse daling van farmaceutische kosten in 2012. Deze kostenontwikkeling was niet meegenomen in het ex-ante vereveningsmodel en daardoor ontstond een verschil tussen de voorspelde en gerealiseerde kosten. Dit betekent een positief, verhogend, effect op het vereveningsresultaat. Als we het effect per individuele verzekerde berekenen, vinden we aanzienlijke verschillen. Kosten van verzekerden die veel medicijnen gebruiken daalden fors terwijl de kosten van gezonde verzekerden niet werden beïnvloed door de daling van farmaceutische kosten.
[15] Dit is een verschil met publicaties waarin naar voorspelbaar winst- of verliesgevende groepen wordt gezocht. Zoals ‘Risicoverevening 2014 voor somatische zorg’ van iBMG.

De beperkte informatie van zorgverzekeraars begrenst de mogelijkheden om het (verwachte) vereveningsresultaat van verzekerden of subgroepen mee te nemen in het bepalen van hun (marketing)strategie.

Daarnaast is het belangrijk dat de signalen die we onderzoeken niet per se duiden op een bewuste actie van zorgverzekeraars om verzekerden te selecteren (Type I acties). Ze geven alleen inzicht in het vereveningsresultaat (in deze fase alleen over de somatische zorg) van de zorgverzekeraar over bepaalde subgroepen van verzekerden. Een positief vereveningsresultaat kan, naast een bewuste actie van de zorgverzekeraar om verzekerden met een laag risico te selecteren, ook op een doelmatige zorginkoop wijzen (hiervoor proberen we te corrigeren in de analyses, zie hoofdstuk 'Indicatoren') of het gevolg zijn van zelfselectie door verzekerden.

In het geval van de eerste onderzoeksvraag (het vereveningsresultaat per subgroep) tonen onze analyses aan dat er imperfecties zitten in het risicovereveningmodel. Dit kan komen door het effect van kenmerken waarmee het model geen rekening houdt of door (kosten)ontwikkelingen die aan de hand van gegevens uit het verleden[16] niet te voorspellen waren. Het eerste effect kan eventueel gecorrigeerd worden door het model te verbeteren, het tweede kan niet eenvoudig binnen het huidige systeem voorkomen worden. Bij de interpretatie van de uitkomsten van de tweede onderzoeksvraag moeten beide effecten wel meegenomen worden.

Een verdere beperking van ons onderzoek is de vraag in hoeverre de resultaten over de periode 2008-2012 te relateren zijn aan de huidige situatie in 2015. Omdat er geen recentere gegevens beschikbaar zijn, kunnen we deze beperking niet oplossen in dit onderzoek. We kunnen wel aanvullende informatie meenemen in het volgende kwalitatieve onderzoek over de huidige situatie en zo een integraal beeld krijgen over de uitgelichte casussen (zie paragraaf 1.4 Fasering van het onderzoek)

1.4 Fasering van het onderzoek

Het onderzoek is zo ingericht dat het jaarlijks herhaald kan worden en de monitoring van de markt continu door kan gaan. Het huidige onderzoek voeren we in verschillende fases uit, omdat we niet over alle benodigde gegevens tegelijk kunnen beschikken. Dit tussenrapport bevat de eerste inzichten. De NZa streeft ernaar om het eindrapport met verdiepende analyses in het laatste kwartaal van dit jaar te presenteren.

In dit tussenrapport focussen we ons op resultaten op het niveau van zorgverzekeraars (risicodragers) en gebruiken we alleen de gegevens van het risicovereveningssysteem die we via het Zorginstituut Nederland krijgen (zie hoofdstuk 3 'Gegevens'). We beperken onze analyses vooralsnog op het somatische model van de verevening. We kijken wel naar een langere periode (2008–2012), zodat we de robuustheid van de resultaten kunnen inschatten. Dat maakt het ook mogelijk naar trends te zoeken.

In de tweede fase van het onderzoek verdiepen we onze analyses en doen we onderzoek op het niveau van afzonderlijke polissen. We breiden ook onze focus uit en kijken naar het integrale resultaat van het model voor de somatische zorg, geestelijke gezondheidszorg en het eigen

---

[16] Het risicovereveningsmodel maakt gebruik van gegevens van drie jaar geleden, want de volledige kostenbestand komt pas drie jaar later beschikbaar.

risico. In dit verdiepende onderzoek kunnen we het werkelijke vereveningsresultaat per verzekerde van de zorgverzekeraar preciezer berekenen en de uitkomsten van de analyses steviger neerzetten. Aan de hand van aanvullende gegevens kunnen we ook een relatie leggen tussen bijvoorbeeld het gemiddelde vereveningsresultaat en de hoogte van het vrijwillig eigen risico of het al dan niet kiezen voor een aanvullende verzekering. De verdiepende analyses betreffen echter een kortere periode, namelijk 2010-2012.

Op basis van de signalen van het kwantitatief onderzoek begint de NZa ook met een kwalitatief onderzoek, waarin individuele casussen worden uitgelicht. Afhankelijk van de uitkomsten van het kwantitatief en kwalitatief onderzoek zal de NZa aanbevelingen doen en zo nodig en mogelijk maatregelen nemen.

1.5 Leeswijzer

In Hoofdstuk 2 geven we een korte beschrijving van het risicovereveningssysteem. Het doel van dit hoofdstuk is om de begrippen in de daaropvolgende analyses toe te lichten. Hoofdstuk 3 beschrijft de gegevensbestanden waarop we de analyses hebben gedaan. Hoofdstuk 4 presenteert de indicatoren en beschrijft hoe deze worden berekend. De resultaten van de analyses zijn in Hoofdstuk 5 beschreven. Hoofdstuk 6 bespreekt deze resultaten en geeft voorlopige conclusies.

2. Het risicovereveningssysteem

Hieronder lichten we de werking van het risicovereveningssysteem op hoofdlijnen toe. Het doel van deze beschrijving is om de begrippen in de daaropvolgende analyses toe te lichten. Hoofdstuk 2.1 'Overzicht vereveningssysteem' is grotendeels overgenomen vanuit de Marktscan Zorgverzekeringsmarkt 2014.[17] De beschrijvingen die we in paragraaf 2.2 'Ex-ante en ex-post in de analyses' gebruiken, zijn gebaseerd op het rapport Beschrijving van het risicovereveningssysteem van de Zorgverzekeringswet[18] terwijl beschrijvingen in paragraaf 2.3 'Tijdschema' op de website van Zorginstituut Nederland opgenomen zijn.[19]

In Nederland zijn zorgverzekeraars verplicht om iedereen te accepteren voor de basisverzekering (acceptatieplicht), ongeacht diens gezondheidskenmerken, ziektehistorie of leefstijl.[20] Het is tegelijk verboden om premiedifferentiatie toe te passen[21] wat in beginsel tot voorspelbare winsten en verliezen leidt door de verschillen in verwachte zorgkosten onder de verzekerden. Risicoverevening is er op gericht om zorgverzekeraars te compenseren voor deze voorspelbare winsten en verliezen die het gevolg zijn van de gezondheidstoestand van hun verzekerden populatie. Verzekerden met een hoog risicoprofiel zijn bijvoorbeeld chronisch zieken, ouderen of verzekerden uit een sociaal-economische risicogroep. Alle zorgverzekeraars krijgen uit het Zorgverzekeringsfonds jaarlijks een financiële bijdrage, die de *vereveningsbijdrage* wordt genoemd. De hoogte van deze bijdrage is afhankelijk van de samenstelling van hun verzekerdenpopulatie.

Als de risicoverevening goed functioneert, heeft de zorgverzekeraar geen financieel gewin bij het weren van verzekerden met hoge verwachte zorgkosten. De risicoverevening zorgt dan voor een gelijke uitgangspositie en daarmee eerlijke concurrentie tussen zorgverzekeraars.

2.1 Overzicht vereveningssysteem

Het risicovereveningssysteem bestaat uit twee onderdelen:
- risicoverevening voorafgaand aan het jaar (ook wel ex-ante verevening genoemd);
- compensaties achteraf (ook wel ex-post compensaties genoemd).

Ex-ante risicoverevening
Om de vergoeding van zorgverzekeraars voor verschillen in het risicoprofiel van hun verzekerden te berekenen, wordt een econometrisch model gebruikt. Dit model voorspelt de zorgkosten aan de hand van persoonskenmerken van verzekerden. Deze voorspelde kosten worden normkosten genoemd. Op basis van de normkosten wordt een vergoeding per zorgverzekeraar berekend. Het model is jaarlijks vastgelegd in de Regeling risicoverevening.

Bij de berekeningen van de normbedragen worden de zorgkosten van verzekerden in meerdere deelbedragen opgedeeld. In dit tussenrapport

[17] Nederlandse Zorgautoriteit, 2014.
[18] Ministerie van Volksgezondheid Welzijn en Sport, 2007.
[19] http://www.zorginstituutnederland.nl/verzekering/risicoverevening+zvw.
[20] Dit is neergelegd in artikel 3, van de Zvw.
[21] Dit is geregeld in artikel 17 van de Zvw.

analyseren we alleen het vereveningsmodel voor somatische zorg[22] en we laten de modellen voor geestelijke gezondheidszorg (ggz), eigen risico en verpleging en verzorging (V&V) buiten beschouwing.[23]

Het model om de normbedragen te berekenen is gebaseerd op de landelijk gemiddelde kosten van verzekerden uitgesplitst naar de risicoprofielen van verzekerden. De normkosten die op basis van het model zijn berekend, zijn dus niet direct afhankelijk van de gerealiseerde kosten van de zorgverzekeraar. Door de jaren heen wordt het vereveningsmodel steeds verbeterd en verfijnd. Er wordt altijd naar nieuwe kenmerken van verzekerden gezocht die de verschillen in zorgkosten nog beter kunnen voorspellen. In Hoofdstuk 3 'Gegevens' staan de gebruikte kenmerken per jaar beschreven.

Ex-post compensaties
De compensaties achteraf worden ingezet voor kosten die moeilijk te voorspellen of door de zorgverzekeraars niet te beïnvloeden zijn. Zo zijn de vaste kostencomponenten in de verevening voor ziekenhuiszorg niet goed te beïnvloeden door zorgverzekeraars en dus worden deze kosten volledig verrekend. Ook kunnen de gegevens die gebruikt worden voor het ex-ante model soms niet representatief worden gemaakt voor een jaar, omdat er bijvoorbeeld grote wijzigingen zijn in de bekostiging van een bepaald segment. Gegevens uit het verleden (waarop het ex-ante risicovereveningsmodel is gebaseerd) en de actuele gegevens sluiten dan niet op elkaar aan. De ex-post compensaties kunnen ook ingezet worden als de kosten door het ex-ante model (tijdelijk) in onvoldoende mate voorspeld worden. In de volgende paragraaf gaan we verder in op de verschillende ex-post compensaties.

2.2 'Ex-ante' en 'ex-post' in de analyses

Zonder risicoverevening is er een aanzienlijke spreiding in het resultaat per verzekerde. Het ex-ante model neemt een groot deel van deze variatie weg. Echter, het gemiddelde resultaat na het ex-ante model is niet noodzakelijk gelijk aan nul wanneer met de werkelijke kosten wordt gerekend. Dit is het gevolg van effecten die niet of achteraf gezien niet juist meegenomen zijn in het ex-ante model, omdat die niet of niet goed te voorspellen waren op het moment van het toepassen van het ex-ante model. Voor deze afwijking op macro niveau corrigeert de macronacalculatie. Het gemiddelde vereveningsresultaat wordt nul na het toepassen van macronacalculatie. Onderlinge verschillen tussen zorgverzekeraars blijven aanwezig.

Omdat wij in onze analyses vereveningsresultaten door de jaren heen willen vergelijken, berekenen we het ex-ante resultaat inclusief de effecten van de macronacalculatie. Voor het gemak noemen we dit het ex-ante resultaat. Deze aanpak zorgt ervoor dat het gemiddelde landelijke resultaat altijd nul is en dus het gemiddelde vereveningsresultaat van een zorgverzekeraar geïnterpreteerd kan worden als afwijking van het landelijk gemiddelde.

---

[22] Onder de kosten van somatische zorg vallen de vaste en variabele ziekenhuiskosten, B-DBC kosten (vanaf 2009) en kosten van overige prestaties (extramurale curatieve zorg zoals de huisartsenzorg). Omdat zorgverzekeraars geen risico dragen voor de vaste ziekenhuiskosten, betrekken we deze niet in onze analyses. De resultaten worden hierdoor niet beïnvloed.
[23] In de tweede fase van dit onderzoek nemen we de modellen voor ggz en eigen risico wel mee. Het model voor V&V is pas vanaf 2015 van toepassing. Omdat we alleen gegevens voor de jaren 2008-2012 hebben, wordt het model V&V niet in ons onderzoek meegenomen.

Na het toepassen van macronacalculatie, worden ook de hoge kosten compensatie en nacalculatie toegepast. Het doel van de hoge kosten compensatie is tegemoet te komen aan kostenverschillen tussen zorgverzekeraars die het gevolg zijn van de ongelijke verdeling van extreem hoge schades over zorgverzekeraars. Nacalculatie wordt ingezet om te compenseren voor tekortschietende beïnvloedingsmogelijkheden van zorgverzekeraars. Vaste ziekenhuiskosten worden bijvoorbeeld 100% verrekend omdat zorgverzekeraars geen invloed hebben op de hoogte van deze kosten.

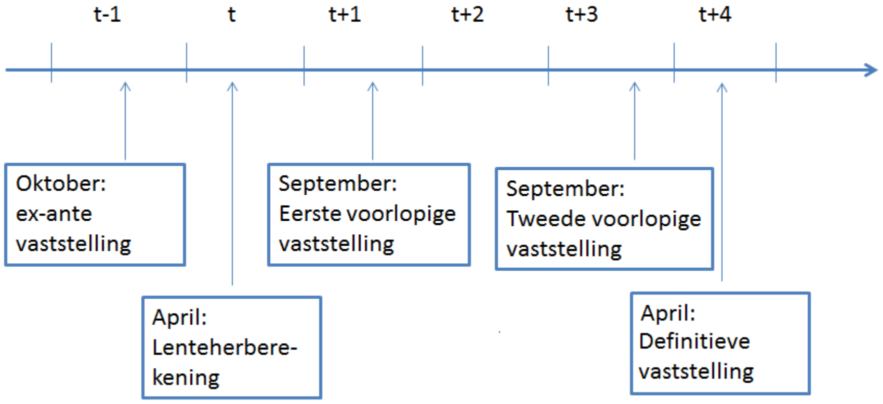
Naast de eerder genoemde ex post compensatiemechanismen kunnen nog tijdelijk aanvullende vangnetten worden gebruikt zoals de bandbreedte regeling. Het doel van vangnetten is het begrenzen van de financiële risico's die zorgverzekeraars tijdelijk lopen. Als het uiteindelijke resultaat (na toepassing van de ex post compensaties) per premiebetalende verzekerde buiten een bepaalde bandbreedte valt, wordt het bedrag buiten de bandbreedte voor bijv. 90% verrekend met het Zorgverzekeringsfonds.

Deze laatste correctie van het vangnet is alleen op het niveau van de zorgverzekeraar (risicodrager) van toepassing. De verschillen in vereveningsresultaat tussen de verzekerden veranderen hierdoor niet en dus blijven de prikkels voor risicoselectie ook onveranderd. Omwille van deze en andere technische redenen genoemd in Bijlage 1 passen we de bandbreedte regeling niet toe in onze analyses. Wanneer wij in dit onderzoek 'ex-post vereveningsresultaat' schrijven, bedoelen we het resultaat na het toepassen van het ex-ante model, de macronacalculatie, de hoge kosten compensatie en de nacalculatie.

2.3 Tijdschema

Het ex-ante model van de risicoverevening wordt in het jaar vóór het gegeven vereveningsjaar berekend. De berekende normbedragen veranderen vervolgens niet meer. Als we het vereveningsjaar 't' noemen, dan wordt de vereveningsbijdrage van elke zorgverzekeraar in oktober t-1 vastgesteld. De bijdragen voor het jaar 2010 werden bijvoorbeeld in oktober 2009 vastgesteld.
Deze vaststelling is echter gebaseerd op voorspelde gegevens. De verzekerdenaantallen en kosten voor het jaar t (bijv. 2010) zijn immers nog niet bekend in t-1 (2009). Om deze redenen is het ook niet mogelijk om ex-post correcties toe te passen. Op het moment dat er meer actuele gegevens beschikbaar komen, wordt de vereveningsbijdrage bijgesteld.

Onderstaande figuur laat de vaststellingsmoment op een tijdslijn zien.



Begin jaar t (2010) worden de verzekerdenaantallen na de overstapperiode bekend. In de lenteherberekening, in april van jaar t (2010) worden daarom de verzekerdenaantallen herberekend.

De kosten van jaar t (2010) uit de verschillende zorgsegmenten zijn pas een aantal jaren later volledig bekend. Daarom zijn er verschillende voorlopige vaststellingsmomenten. In september van jaar t+1 (2011) vindt de eerste voorlopige vaststelling plaats. Hierbij wordt de vereveningsbijdrage per zorgverzekeraars op basis van gerealiseerde verzekerdenaantallen en voorlopige gegevens in de jaarstaat Zorgverzekeringswet (Zvw) van de zorgverzekeraars voorlopig vastgesteld. Met uitzondering van de hoge kosten compensatie, zijn ook de ex-post compensaties toegepast. In september van jaar t+3 (2013) wordt de vereveningsbijdrage per zorgverzekeraar voor de tweede keer voorlopig vastgesteld (tweede voorlopige vaststelling) op basis van gedeeltelijk nieuwe gegevens. Verder betrekt het Zorginstituut hierbij de opbrengstverrekening van ziekenhuizen, de opgave hoge kosten compensatie van de zorgverzekeraars en de dan bekende correcties van de NZa op basis van de rapporten Zvw. De definitieve vaststelling vindt uiterlijk in jaar t+4 (2014) plaats. Hierbij worden eventuele correcties doorgevoerd op de definitieve gegevens in de jaarstaat Zvw op basis van de rapporten Zvw van de NZa.

3. Gegevens

De data voor deze eerste fase van het onderzoek naar risicoselectie verkregen we van het Zorginstituut Nederland. We maken gebruik van bestanden van verzekerdenkenmerken, werkelijke zorgkosten en normbedragen. De gegevens zijn op gepseudonimiseerd BSN-niveau beschikbaar.[24] De rest van dit hoofdstuk bestaat uit een beknopte beschrijving van de data die relevant zijn voor het somatische model en een toelichting op de discrepanties tussen onze dataset en de gegevens gebruikt in de risicoverevening door het Zorginstituut.

3.1 Omschrijving bestanden

We maken gebruik van gegevens op het niveau van verzekerden over de periode 2008-2012. Dit zijn gegevens over kenmerken en werkelijk gemaakte kosten van verzekerden die ten grondslag liggen van de vaststellingen bij de risicoverevening voor somatische zorg. De kostengegevens zijn gebaseerd op de 'kosten per verzekerde'-bestanden. Het betreft altijd de gedeclareerde kosten tot en met 31 december twee jaar na het vereveningsjaar (12 kwartalen), behalve het bestand van 2012 (ongeveer 9 kwartalen). De verzekerdenkenmerken zijn gegevens van de meest recente vaststelling.[25]

Tabel 3.1 toont de beschrijving van variabelen in de verzekerdenbestanden die we van Zorginstituut Nederland hebben verkregen. In het verzekerdenbestand van elk jaar zitten alle kenmerken van verzekerden die in de vereveningsmodellen van dat jaar gebruikt zijn. Deze kenmerken bepalen de indeling van verzekerden in risicoklassen. Elke waarneming in het bestand wordt geïdentificeerd aan de hand van een gepseudonimiseerd burgerservicenummer van de verzekerede en een uzovi-nummer van de zorgverzekeraar. Met deze twee variabelen valt het verzekerdenbestand te koppelen aan het kostenbestand. De variabele 'Weging' geeft aan hoe lang de verzekerde ingeschreven stond bij de betreffende zorgverzekeraar. Deze variabele gebruiken we in de analyses om de 'verzekerdenjaren' te berekenen. Als een baby bijvoorbeeld op 1 juli geboren wordt, staat hij maar voor een half jaar ingeschreven bij de zorgverzekeraar. Hij telt mee in de populatie van de zorgverzekeraar met een half verzekerdenjaar. Voor verzekerden die midden in het jaar van één naar een andere zorgverzekeraar zijn overgestapt, zijn er meerdere records in het bestand. In de analyses rekenen we altijd met verzekerdenjaren. In het vervolg bedoelen we 'resultaat per verzekerdenjaar' wanneer we 'resultaat per verzekerde' schrijven.

[24] Omdat wij voor meerdere indicatoren de verzekerden moeten identificeren wie van een verzekeraar naar de andere overstappen, hebben we de gegevens op verzekerdenniveau nodig. In het bestand staan persoonsgegevens gepseudonimiseerd waardoor de gegevens niet meer tot persoon herleidbaar zijn.

[25] Voor het jaar 2008 is het verzekerdenbestand van de tweede voorlopige vaststelling gebruikt. Alleen bij SES zijn er opmerkelijke verschillen ten opzichte van de definitieve vaststelling. Dit verschil heeft marginaal effect op ons resultaat en verandert onze conclusies niet.

Tabel 3.1. Variabelen in het verzekerdenbestand

Variabele	Jaren	Beschrijving
BSN	2008 – 2012	Gepseudonimiseerd burgerservicenummer
UZOVI	2008 – 2012	Identificatienummer zorgverzekeraar
Postcode	2008 – 2012	De vier cijfers van de Nederlandse postcode
Weging	2008 – 2012	De fractie van het jaar dat een verzekerde ingeschreven is. Dit is vooral van belang voor personen die 18 jaar worden en tijdens het jaar overstappen van zorgverzekeraar.
Leeftijd en geslacht	2008 – 2012	Leeftijdsgroep (per 5 jaar) naar geslacht. Tot met 2009 38 groepen, vanaf 2010 40 groepen (0-jarigen zijn als aparte groep opgenomen). In 2008 was de knip tussen leeftijdsgroepen bij 20 jarigen, vanaf 2009 was deze knip verlaagd naar 18 jarigen.
DKG	2008 – 2012	Diagnosekostengroepen gebaseerd op de diagnoses die verzekerden krijgen bij hun ontslag uit het ziekenhuis die waarschijnlijk hoge toekomstige zorgkosten met zich mee brengen. In totaal 14 groepen inclusief de groep zonder DKG.
FKG	2008 – 2012	Farmaceutische kostengroepen, ter compensatie van hoge kosten door het gebruik van extramuraal afgeleverde geneesmiddelen. De kostengroepen veranderen door de tijd met 20 groepen in 2008 en 2009, 23 groepen in 2010 en 2011 en 25 groepen in 2012, inclusief de groep zonder medicijngebruik.
APE	2008 – 2012	Regiokenmerk somatische zorg met in totaal 10 regio's op basis van een clustering van postcodegebieden.
AVI	2008 – 2012	Aard van inkomen met in alle jaren een uitsplitsing tussen arbeidsongeschikten, bijstandsontvangers, zelfstandigen en de referentiegroep naar vier leeftijdsgroepen (16 groepen) met daarnaast kinderen en ouderen (2 groepen). In 2008 en 2009 worden WW'ers en overige uitkeringsontvangers ook als aparte groep aangemerkt (4 groepen).
SES	2009 – 2012	Sociaal economische status op basis van de hoogte van het fiscaal inkomen. In totaal 3 inkomensgroepen (inclusief referentiegroep) onderverdeeld in 3 leeftijdsgroepen (9 groepen) en een groep mensen woonachtig op adressen met veel bewoners naar leeftijd (3 groepen)
MHK	2012	Meerjarige hoge kosten over de afgelopen 3 jaren. In totaal 7 groepen

Tabel 3.2 geeft een overzicht van de variabelen in het kostenbestand die we van het Zorginstituut Nederland hebben verkregen. De afzonderlijke records worden geïdentificeerd door gepseudonimiseerde burgerservicenummers van verzekereden en uzovi-nummers van hun zorgverzekeraars. Verder zijn er drie kostenvariabelen: variabele kosten medisch specialistische zorg (MSZ), schade in het B-DBC segment en schade aan overige zorgkosten die door zorgverzekeraar worden betaald. Het aantal records in het kostenbestand is kleiner dan het aantal records in het verzekerdenbestand, want verzekerden die geen zorg hebben gebruikt in een bepaald jaar kunnen in de kostengegevens ontbreken. Het verschil tussen aantallen in de twee bestanden is in elk jaar

maximaal 2% van alle records. In het geval dat een verzekerde niet in het kostenbestand zit, veronderstellen we dat zijn kosten nul zijn.

Tabel 3.2. Variabelen in het kostenbestand*

Variabele	Jaren	Beschrijving
BSN	2008 – 2012	Gepseudonimiseerd burgerservicenummer
UZOVI	2008 – 2012	Identificatienummer zorgverzekeraar
Variabel MSZ	2008 – 2012	Variabele kosten gebaseerd op de DBC (tot en met 2011) en DOT systematiek (vanaf 2012) inclusief bijvoorbeeld kosten van vrijgevestigde specialisten en specialisten in loondienst, maar exclusief de kosten van zorgproducten in het vrije segment (B-DBC's)
Schade B-DBC's	2009 – 2012	Kosten van alle DBC's in het B-segment
Schade overige	2008 – 2012	Kosten van alle overige prestaties (bijvoorbeeld farmaceutische zorg en huisartsenzorg) die niet geclassificeerd zijn als medisch specialistisch zorg.

* Het deelbedrag vaste ziekenhuiskosten zit niet in ons kostenbestand. Omdat deze kosten volledig verrekend worden, beïnvloeden deze niet de resultaten. Daarom worden vaste ziekenhuiskosten ook in onze analyse niet meegenomen.

In onze analyses gebruiken we twee sets van normbedragen; beide afkomstig van het Zorginstituut Nederland. Beide sets bestaan uit de normbedragen per risicoklasse voor drie typen kosten (variabele kosten van medisch specialistische zorg, kosten B-DBC en overige zorgkosten).

De eerste set van normbedragen van Zorginstituut Nederland die we hier gebruiken, heet 'ex-post normbedragen voor het kostenbudget'. Dit zijn de normbedragen waarop alleen de macronacalculatie is toegepast, en geen hoge-kosten-correcties (HKC) en nacalculaties van vereveningsresultaten zijn toegepast. De tweede set van normbedragen heet 'ex-post normbedragen voor afrekenbudget'. Dit zijn de normbedragen waarop HKC en nacalculaties wel zijn toegepast. Met de eerste set van normbedragen worden onze 'ex-ante vereveningsresultaten' berekend, en met de tweede set worden onze 'ex-post vereveningsresultaten' berekend. Hierna gebruiken we daarom de termen 'ex-ante normbedragen' en 'ex-post normbedragen' om naar de respectievelijke sets van normbedragen te verwijzen.

3.2 Analysebestand vs. Risicoverevening-gegevens

Er is een kleine discrepantie tussen onze dataset en de volledige dataset van het Zorginstituut Nederland die voor de definitieve en voorlopige vaststellingen zijn gebruikt waardoor we niet alle gegevens hebben die voor de vaststelling nodig zijn.

Ten eerste bevat ons verzekerdenbestand alleen verzekerden van wie oorspronkelijk een BSN-nummer bekend was, terwijl de definitieve vaststelling van het Zorginstituut ook rekening houdt met verzekerden met een foutief of ontbrekend BSN-nummer. Deze discrepantie heeft echter geen materiele gevolgen voor onze conclusies omdat het effect op de resultaten verwaarloosbaar is.[26]

Ten tweede beperken we onze analyse tot gegevens van verzekerden woonachtig in Nederland geïdentificeerd aan de hand van de variabele APE (regiokenmerk, zie ook in Tabel 3.1). De reden voor deze keuze is

[26] Zo'n 500 verzekerdenjaren in 2008.

dat de curatieve zorgkosten gemaakt in het buitenland niet in onze kostenbestanden zitten. Door verzekerden die in het buitenland wonen, uit te sluiten, corrigeren we in een bepaalde mate voor het ontbreken van deze kosten. De interpretatie van onze resultaten is daarom ook beperkt tot de in Nederland woonachtige verzekerden. Het aantal verzekerden dat in het buitenland woont, betreft doorgaans een klein percentage van de totale populatie per zorgverzekeraar. Het ging in 2008 bijvoorbeeld om minder dan 1,5% van de verzekerdenjaren bij alle zorgverzekeraars met uitzondering van twee risicodragers die hogere percentages hadden.

Verder zijn er ook kleine discrepanties met de vaststelling van het Zorginstituut door enkele kostenposten die niet in onze bestanden zitten. Dit betreft:

- nagekomen declaraties en creditnota's ouder dan twee jaar voor het jaar waarover de analyse gaat (in de vaststelling van het Zorginstituut 2009 zijn bijvoorbeeld ook kosten uit 2008 en eerder meegenomen);
- kosten gemaakt ten behoeve van de stichting kwaliteitsgelden medisch specialisten (SKMS, een vast bedrag per verzekerde);
- kosten ter ondersteuning van de eerstelijnszorg;
- De opbrengstverrekening is in het kostenbestand op verzekerdenniveau gebeurd. Bij de vaststelling door het Zorginstituut is de opbrengstverrekening verwerkt op het niveau van zorgverzekeraars (risicodragers). Beide bronnen sluit niet exact aan.

4. Indicatoren

Dit hoofdstuk beschrijft de indicatoren die we gebruiken om onze onderzoeksvragen over subgroepen (paragraaf 4.1) en selectieve in- en uitstroom (paragraaf 4.2) te beantwoorden. Alle indicatoren in dit onderzoek zijn uitgedrukt in een gemiddeld vereveningsresultaat per verzekerdenjaar over een bepaalde populatie van verzekerden. Dit kan zowel een subgroep van verzekerden als de hele populatie van een zorgverzekeraar zijn. In alle analyses maken we onderscheid tussen het ex-ante en het ex-post vereveningsresultaat (voor de definitie zie Hoofdstuk 2).

4.1 Vereveningsresultaat per subgroep

Om een idee te krijgen of met bepaalde subgroepen van verzekerden risicoselectie voorkomt, berekenen we het vereveningsresultaat van *subgroep* g in jaar t als gemiddelde waarde van het vereveningsresultaat van alle verzekerden in deze subgroep over verzekerden in Nederland[27]:

(1) $$S_{g,t} = \frac{\sum_{i \in subgroep\ g}(b_t X_{it} - C_{it})}{N_{subgroep\ g,t}}$$

N_{gt} is het aantal verzekerden in subgroep g; b_t zijn de normbedragen in het jaar t; X_{it} zijn de respectievelijke kenmerken van verzekerde i waarvoor het risicovereveningssysteem (verder RV systeem) in het jaar t compenseert (zoals leeftijds- en geslachtklassen, regioklassen, enz.); C_{it} staat voor de werkelijke kosten van verzekerde i in het jaar t.

In deze eerste fase van het onderzoek identificeren we alleen groepen op basis van de vereveningskenmerken We berekenen het werkelijke vereveningsresultaat voor bijvoorbeeld de subgroepen van 65-plussers en jongeren. Hoewel het vereveningsmodel voor leeftijdskenmerken corrigeert, is het mogelijk dat bepaalde subgroepen in bepaalde jaren meer of minder winstgevend zijn.[28]

4.2 Vereveningsresultaat per zorgverzekeraar

De tweede onderzoeksvraag beantwoorden we aan de hand van vijf indicatoren die *per zorgverzekeraar* per jaar worden uitgerekend. De indicatoren geven het volgende weer:

- Het gemiddelde vereveningsresultaat op de totale populatie van de zorgverzekeraar (indicator T).
- Het gemiddelde vereveningsresultaat op de uitstroom, dus op de verzekerden die overstappen naar een andere zorgverzekeraar (indicator U).

---

[27] Om de redenering in dit hoofdstuk simpel te houden, presenteren we formules onder de aanname dat elke verzekerde het hele jaar bij dezelfde zorgverzekeraar zit. De berekening die daarna plaats vindt houdt, rekeningen met gewichten van verzekerden in het verzekerdenbestand van de verzekeraar. Zie meer detail hierover in Bijlage 2.

[28] Dit komt door mogelijke onverwachte ontwikkelingen tussen het jaar *t*-3 (waarmee de risicoverevening berekening wordt gedaan) en het jaar *t* (waarvoor de uitkomsten van het risicoverevening model zijn toegepast).

- Het gemiddelde vereveningsresultaat op de instroom, dus voor de nieuwe verzekerden in een bepaald jaar berekend vóór en ná de overstap (indicatoren I^* en I).
- Het effect van niet voorspelbare kostenontwikkelingen op het resultaat van de zorgverzekeraar gezien de samenstelling van zijn populatie (indicator V).

4.2.1 Analyse van de totale populatie

Het vereveningsresultaat op de totale populatie van zorgverzekeraar k in het jaar t ($T_{k,t}$) is het verschil tussen de gemiddelde zorgkosten die op basis van de normbedragen zijn berekend en de werkelijke zorgkosten per verzekerde in de totale populatie van zorgverzekeraar k in het jaar t:

$$T_{k,t} = \frac{\sum_{i \in populatie\ k}(b_t\, X_{it} \ - C_{it})}{N_{k,t}} \qquad (2)$$

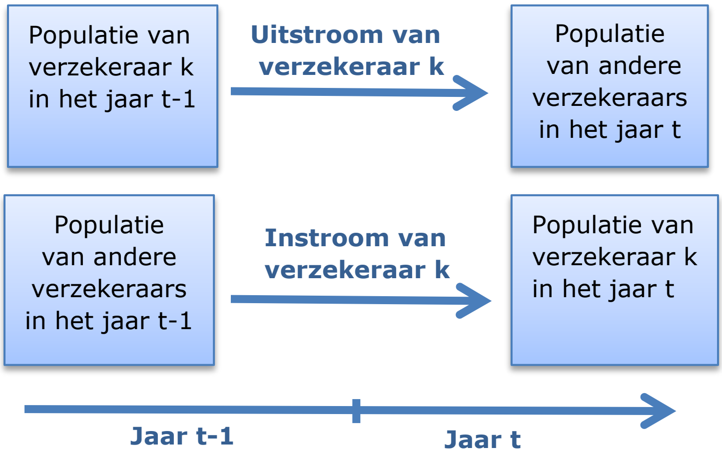
Hierbij gebruiken we dezelfde notatie als in paragraaf 4.1.Deze indicator laat zien in hoeverre de werkelijke kosten van zorgverzekeraar k in het jaar t via het RV systeem worden gecompenseerd. We analyseren de spreiding van deze indicator over de risicodragers en onderzoeken onder andere hoe deze indicator over de periode 2008-2012 verandert. Ook bekijken we hoe het komt dat steeds dezelfde zorgverzekeraars gunstige of ongunstige uitkomsten behalen.

Het is belangrijk om op te merken dat een positief of negatief vereveningsresultaat op de totale populatie ($T_{k,t}$) op zichzelf weinig over risicoselectie zegt. Het resultaat kan namelijk ook door verschillen in doelmatigheid (bijv. scherpere inkoopprijzen) tot lagere kosten leiden. Het is ook mogelijk dat een zorgverzekeraar toevallig in bepaalde jaren een meer winstgevende populatie had. De indicator T die alleen de totale eigen kosten van een zorgverzekeraar weergeeft, kan de andere effecten niet onderscheiden.

4.2.2 Analyse van in- en uitstroom

Om de verschillende effecten van elkaar te onderscheiden, kan de analyse van in- en uitstroom van de zorgverzekeraars een oplossing bieden. We beschouwen een verzekerde als een ‘overstapper’; als hij in jaar t bij een andere zorgverzekeraar zit dan in jaar t-1. Voor de zorgverzekeraar waar hij in jaar t-1 was ingeschreven, maakt hij deel uit van de uitstroom, terwijl hij voor de zorgverzekeraar in jaar t is opgenomen in de instroom. De in- en uitstroom is gedefinieerd op basis van de overstappers. Geboorte en overlijden rekenen we niet mee in de in- of uitstroom. Figuur 4.1 hieronder toont een schematisch beeld van de in- en uitstroom van zorgverzekeraar k in het jaar t. In Bijlage 2 ‘Technische toelichting’ staat in detail opgeschreven hoe we overstappers en in- en uitstroom van elke zorgverzekeraar identificeren.

Figuur 4.1. Instroom en uitstroom in het jaar t



De indicatoren U, I en I* focussen op overstappers: ze betreffen de in- en uitstroom uit de populatie van een zorgverzekeraar in het jaar van de overstap, of net daarvóór (wanneer deze verzekerden nog bij andere zorgverzekeraars zaten).

Het gemiddelde vereveningsresultaat voor de uitstroom van verzekerden, $U_{k,t}$, is het gemiddelde verschil tussen de kosten die onder het RV systeem worden berekend en de werkelijke kosten van verzekerden in de uitstroom van zorgverzekeraar k in het jaar t:

(3) $$U_{k,t} = \frac{\sum_{i \in uitstroom\, k,t}(b_t X_{it} - C_{it})}{N_{uitstroom\, k,t}}$$

Hierbij nemen we de uitstroom die net vóór het jaar t plaats vindt (zie Figuur 1); we berekenen het vereveningsresultaat voor deze groep in het jaar t. Aangezien de uitstroom naar meerdere zorgverzekeraars gaat, worden de kosten van deze groep van verzekerden uitgemiddeld.[29] Zo corrigeren we voor het effect van doelmatigheidsverschillen tussen zorgverzekeraars op het vereveningsresultaat. De indicator geeft daarmee aan of de uitstroom van zorgverzekeraar k een relatief gunstig of ongunstig risicoprofiel heeft.

Een aanhoudend negatief resultaat voor de uitstroom bij een zorgverzekeraar in combinatie met een positief vereveningsresultaat voor de gehele populatie geeft ons een signaal van mogelijke risicoselectie. Op zichzelf is dit geen definitief oordeel, want andere indicatoren kunnen dit signaal zowel afzwakken als versterken. Het kan bijvoorbeeld zo zijn dat het vereveningsresultaat over de instroom (zoals beschreven hieronder) van dezelfde zorgverzekeraar aanhoudend positief is. In dit geval versterken beide signalen elkaar en krijgen ze meer gewicht in de conclusie.

Het gemiddelde vereveningsresultaat over de instroom van verzekerden ($I_{k,t}$) is het gemiddelde verschil tussen de kosten die onder het RV-

[29] Dit geld onder de aanname dat de verzekerden met ongeveer dezelfde kans bij de andere zorgverzekeraars instromen. Mocht een groot deel van de uitstroom bestaan uit één grote collectiviteit die in zijn geheel naar een andere verzekeraar instroomt, dan geldt deze aanname niet. In dat bijzondere geval kunnen we niet voor doelmatigheid corrigeren.

systeem zijn berekend en de werkelijke kosten van verzekerden in de instroom van verzekerden van zorgverzekeraar k. Deze indicator berekenen we voor het jaar t en ook voor het jaar t-1:

$$I_{k,t} = \frac{\sum_{i \in instroom\ k,t} (b_t\ X_{it} - C_{it})}{N_{instroom\ k,t}} \quad (4)$$

$$I^*_{k,t} = \frac{\sum_{i \in instroom\ k,t} (b_{t-1}\ X_{i,t-1} - C_{i,t-1})}{N_{instroom\ k,t}} \quad (5)$$

De waarde van de indicator in jaar t laat het resultaat van zorgverzekeraar k zien op de instroom. Deze waarde wordt vergeleken met de waarde in het jaar (t-1) toen deze verzekerden nog bij andere zorgverzekeraars zaten. Bij eventuele risicoselectie verwachten we een structureel positief resultaat voor de instroom door de jaren heen. Als de instroom echter door doelmatigheid van de zorgverzekeraar positief is, verwachten we een resultaat van I* rond nul. Hierbij is het idee hetzelfde als bij het berekenen van de uitstroom. Omdat verzekerden in jaar t-1 bij verschillende zorgverzekeraars verzekerd zijn middelen eventuele doelmatigheidsverschillen uit. Door het vergelijken van het vereveningsresultaat van de instroom van jaar t in het jaar t met t-1, kunnen we grotendeels voor het effect van doelmatigheid in het vereveningsresultaat over de instroom corrigeren.[30]

Wanneer de waarde van indicator I negatief en de waarde van indicator I* positief is, is ondoelmatigheid van de zorgverzekeraar niet de enige logische verklaring. We verwachte deze resultaten ook wanneer verzekerden een nieuwe zorgverzekeraar kiezen omdat zij hogere zorgkosten verwachten. In jaar t-1 hebben deze verzekerden lage of gemiddelde kosten maar in het jaar daarop wanneer ze bij de nieuwe zorgverzekeraar instromen, maken ze hoge kosten. In dit geval is het vereveningsresultaat over de instroom lager in jaar t dan in jaar t-1.

4.3 Effect van niet voorspelbare kostenontwikkelingen

De laatste indicator – indicator V – is een maat voor het effect van niet voorspelbare kostenontwikkelingen en aannames van het ex-ante vereveningsmodel die achteraf niet juist blijken te zijn op het resultaat van de zorgverzekeraar afhankelijk van zijn specifieke verzekerdenpopulatie. Hiervoor berekenen we eerst een normbedrag op basis van de werkelijke kosten van jaar t. Dit nieuwe normbedrag bevat alle beschikbare informatie. Dit in tegenstelling tot het werkelijke normbedrag dat op de verwachte kosten gebaseerd is. Het verschil tussen de opnieuw geschatte en de werkelijke normbedragen is het effect van het verschil tussen de aannames van het ex-ante model en wat in het werkelijkheid gebeurde. Door de specifieke samenstelling van de verzekerdenpopulatie van elke zorgverzekeraar kan het totale effect per zorgverzekeraar verschillen. De indicator V laat dus zien hoe niet

[30] Een aandachtpunt hierbij is dat het verschil tussen I en I* meer is dan doelmatigheid alleen, omdat de normbedragen in deze twee jaar (b_t en b_{t-1}) ook verschillen. Hierdoor zijn ook andere effecten meegenomen. Verder geld hierbij dezelfde uitzondering over grote collectiviteiten als bij de uitstroom (zie voetnoot 29).

voorspelbare (kosten)ontwikkelingen het resultaat van de specifieke zorgverzekeraar hebben beïnvloed.

Deze indicator wordt als volgt berekend. Eerst berekenen we het RV-model voor het jaar t op basis van de kosten en kenmerken van hetzelfde jaar:

$$C_{it} = \alpha_t X_{it} + \varepsilon_{it}$$

Dit model wordt geschat met de definitieve verzekerdenkenmerken en de werkelijke kosten van jaar t. De geschatte coëfficiënten a_t noemen we 'opnieuw geschatte normbedragen'. De notatie ε_{it} staat voor een storingsterm. De voorspelde kosten uit dit model zijn gelijk aan de gemiddelde kosten van verzekerden (afhankelijk van hun kenmerken).

De opnieuw geschatte normbedragen zijn de 'ideale' normbedragen van het RV systeem voor het jaar t als de daadwerkelijk gerealiseerde kosten en kenmerken van tevoren bekend zouden zijn. De opnieuw geschatte (a_t) normbedragen kunnen zowel ex-ante als ex-post afwijken van de werkelijke normbedragen (b_t). Het verschil (b_t – a_t) geeft weer hoe goed het RV systeem van het jaar t het effect van elk kenmerk op de kosten van tevoren kon inschatten. Beleidswijzigingen waarmee geen rekening werd gehouden in het ex-ante model of andere veronderstellingen van het model die achteraf niet juist blijken, resulteren in een verschil tussen de werkelijke en opnieuw geschatte normbedragen.

Met de indicator V bekijken we wat is het effect van dit verschil (b_t – a_t) per kenmerk op het niveau van de zorgverzekeraar (risicodrager) en zijn specifieke verzekerdenpopulatie. De verwachting is dat het effect groter is bij kleine zorgverzekeraars omdat zij minder verzekerden hebben om het risico van niet voorspelbare kostenontwikkelingen te dekken. De indicator V wordt als volgt berekend:

$$(6) \qquad V_{k,t} = \frac{\sum_{i \in populatie\ k} (b_t - a_t)\, X_{it}}{N_{k,t}}$$

Op basis van deze formule kan het verschil tussen het werkelijke resultaat van verzekerde i en het effect van de verschuiving van de normbedragen als volg uitgedrukt:

$$(b_t X_{it} - C_{it}) - (b_t - a_t) X_{it} = a_t X_{it} - C_{it} = -\hat{\varepsilon}_{it}$$

De storingsterm bevat de afwijkingen van het gemiddelde werkelijke resultaat. Op het niveau van een verzekeraar betekent dit in formulevorm het volgende:

$$(7) \qquad T_{k,t} - V_{k,t} = -\frac{\sum_{i \in populatie\ k} \hat{\varepsilon}_{it}}{N_{k,t}}$$

Op deze manier kan het werkelijke resultaat worden 'geschoond' van de onvoorspelbare effecten op normbedragen die voor alle zorgverzekeraars een rol spellen. Wat overblijft is de afwijking van het RV-resultaat per verzekerde van verzekeraar k van het RV-resultaat per verzekerde in Nederland. Deze afwijking kan nog steeds zowel het effect van de doelmatigheid als het effect van risicoselectie bevatten, maar is

geschoond van de effecten die door de verschuiving van normbedragen zelf komen. Bij de interpretatie van resultaten is het daarom ook interessant om naar het verschil tussen T en V te kijken.

5. Resultaten

In dit hoofdstuk geven we de resultaten van onze analyses weer. In paragraaf 5.1 bespreken we de resultaten over de subgroepen. Paragraaf 5.2 laat de uitkomsten van indicatoren T, U en I zien, en paragraaf 5.3 is gewijd aan indicator V. In paragraaf 5.4 geven we een overzicht van alle indicatoren per zorgverzekeraar en selecteren we de meest interessante gevallen. De resultaten van de verschillende indicatoren voor de geselecteerde zorgverzekeraars analyseren we in samenhang in paragraaf 5.5.

Alle berekeningen in dit hoofdstuk zijn gedaan voor de dataset van de Nederlandse bevolking. Deze dataset komt overeen met de dataset van de vaststellingen van Zorginstituut Nederland, met enkele kleine verschillen, namelijk: (1) onze set bevat alleen de verzekerden met een geldig BSN en woonachtig in Nederland; (2) aan de kostenzijde zijn enkele kostenposten niet meegenomen (zie hoofdstuk 'Gegevens' voor details). Omdat we ons alleen focussen op de opmerkelijke verschillen tussen de relatieve uitkomsten van zorgverzekeraars, kunnen deze kleine verschillen de conclusies niet beïnvloeden.

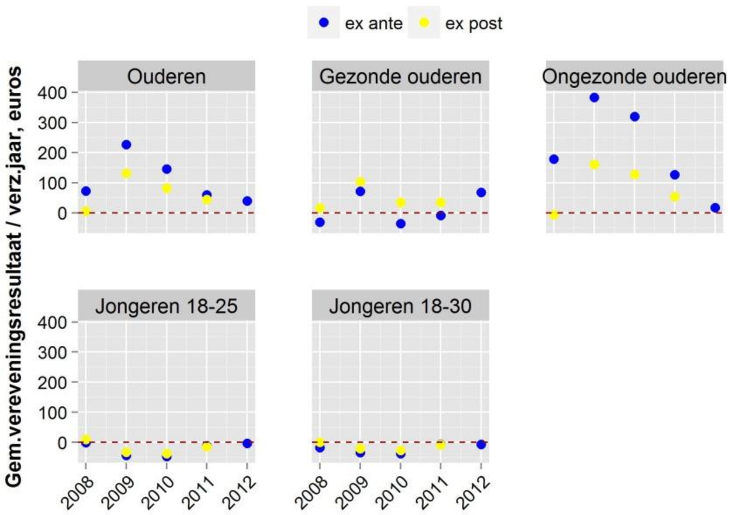
5.1 Resultaten per subgroep

In dit tussenrapport bekijken we de groepen 'ouderen' en 'jongeren'. Ouderen maken doorgaans meer zorgkosten dan jongeren. Het risicovereveningssysteem corrigeert voor deze verschillen. Zonder het vereveningssysteem zouden ouderen 'verliesgevend' en jongeren 'winstgevend' zijn voor de zorgverzekeraars. Door de compensaties via het risicovereveningssysteem wordt verwacht dat het gemiddelde vereveningsresultaat van jongeren en ouderen op 0 uitkomt. Door de werkelijke kosten te gebruiken in de berekeningen, kunnen we vaststellen wat het gemiddelde vereveningsresultaat is voor deze subgroepen in de periode 2008-2012.

Figuur 5.1 laat het vereveningsresultaat per verzekerde zien voor de subgroepen ouderen (65 jaar en ouder) en jongeren. Als gevoeligheidstoets gebruiken we twee verschillende definities van de 'jongeren' (18-25 jaar en 18-30 jaar). De subgroep 'ouderen' splitsen we in twee kleinere subgroepen, namelijk:

- 'gezonde ouderen'; degenen die in geen FKG, DKG of MHK-klasse zijn ingedeeld.
- 'ongezonde ouderen'; degenen die wel in een FKG, DKG of MHK-klasse zijn ingedeeld.

Figuur 5.1. Vereveningsresultaat per subgroep*



Definitie subgroep Gezonde ouderen: leeftijd 65+ zonder FKG, DKG of MHK in 2012
Definitie subgroep Ongezonde ouderen: leeftijd 65+ met FKG of DKG of MHK in 2012
Andere splitsing naar leeftijdsgroepen in 2008: Jongeren vanaf 15 jaar

* in 2012 werden de ex-post compensaties afgeschaft waardoor geen ex-post resultaat meer bestaat

Uit Figuur 5.1 blijkt dat de groep 'ouderen' in 2009 een hoog positief gemiddeld vereveningsresultaat had in het model voor somatische zorg, dat in de jaren daarop geleidelijk afnam. We zien hetzelfde patroon maar met nog hogere positieve resultaten voor de groep ongezonde ouderen. Het lijkt erop dat dit patroon van resultaten voornamelijk is bepaald door een aantal DKG-risicoklassen die in de betreffende jaren overgecompenseerd waren in de verevening. De subgroep 'gezonde ouderen' had een lager, maar nog steeds positief gemiddeld vereveningsresultaat dat in 2010 en 2011 verder daalde. Het ex-ante resultaat werd zelfs negatief.

Het resultaat van de groep 'jongeren' is licht negatief in het model voor somatische zorg in 2009 en 2010 ongeacht de precieze leeftijdsgrens. De voorlopige conclusie op basis van de resultaten van het model voor somatische zorg is, dat een zorgverzekeraar die veel jongeren en weinig ouderen onder zijn verzekerden had, slechter af was dan een zorgverzekeraar met veel ouderen en weinig jongeren in zijn populatie.

5.2 Resultaten per zorgverzekeraar over T, U, I, I*

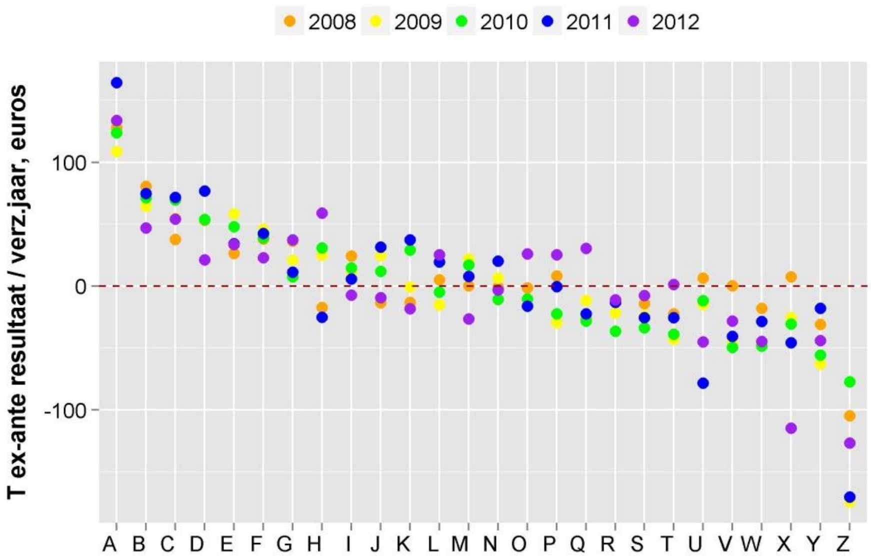
De volgende figuren laten de gemiddelde resultaten zien per verzekerde van alle zorgverzekeraars voor de indicatoren T, U en I. Elk figuur focust op één indicator uit het hoofdstuk 4 'Indicatoren'. De zorgverzekeraars zijn in alle figuren geordend naar de hoogte van hun gemiddelde vereveningsresultaat per verzekerde over de totale populatie (van hoog naar laag) en worden geanonimiseerd weergegeven. Dezelfde letter wijst altijd naar dezelfde zorgverzekeraar in alle figuren.

In deze paragraaf tonen we alleen de ex-ante vereveningsresultaten. De macronacalculatie is er wel in verwerkt (voor 2012 zijn soortgelijke effecten gesimuleerd) zodat de verschillende jaren vergelijkbaar zijn. Voor de jaren 2008-2011 geeft deze analyse een beeld hoe het vereveningsresultaat van de zorgverzekeraars was geweest zonder

verdere ex-post compensaties. In Bijlage 1 vindt u de figuren met de ex-post resultaten.

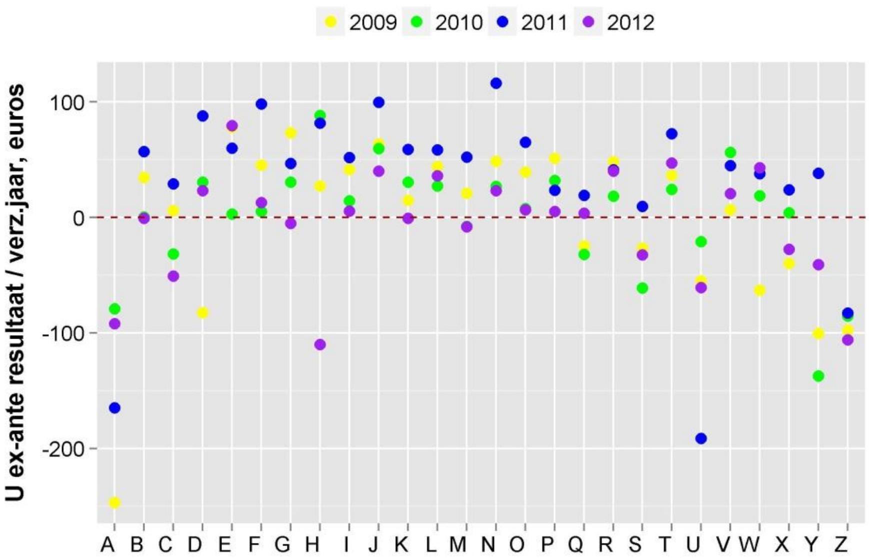
Figuur 5.2 laat zien dat de resultaten van ongeveer de helft van de zorgverzekeraars op hun totale populatie in een range van -50 en +50 euro schommelen. Deze zorgverzekeraars hadden in sommige jaren in de somatische zorg hogere en in andere jaren lagere gemiddelde zorguitgaven per verzekerde dan door het vereveningsmodel voorspeld werd. Hun resultaten lijken dus door eenmalige gebeurtenissen bepaald te worden. Er is een aantal zorgverzekeraars die structureel in elk onderzocht jaar een positief of negatief vereveningsresultaat behaalden. Er zijn ook een paar opvallende uitschieters die in meerdere jaren een hoger dan +50 euro of lager dan -50 euro vereveningsresultaat per verzekerde behaalden in het model voor de somatische zorg. In deze laatste twee gevallen zijn er aanwijzingen dat de verzekerdenpopulatie van deze zorgverzekeraars structureel van het gemiddelde afwijkt (een mogelijke signaal van risicoselectie) of dat hun zorguitgaven per verzekerde structureel van het gemiddelde afwijken (een mogelijke signaal voor doelmatigheid).

Figuur 5.2. Vereveningsresultaat per verzekerde op de totale populatie van een zorgverzekeraar



Figuur 5.3 laat het gemiddelde vereveningsresultaat van uitstromers per verzekerde zien voor elke zorgverzekeraar. Het gemiddelde resultaat voor overstappers is positief zoals in de figuur ook opvalt. Dit betekent dat verzekerden die overstappen doorgaans lagere zorguitgaven hebben dan door het vereveningsmodel voorspeld. Overstappers hebben dus een gunstig risicoprofiel.

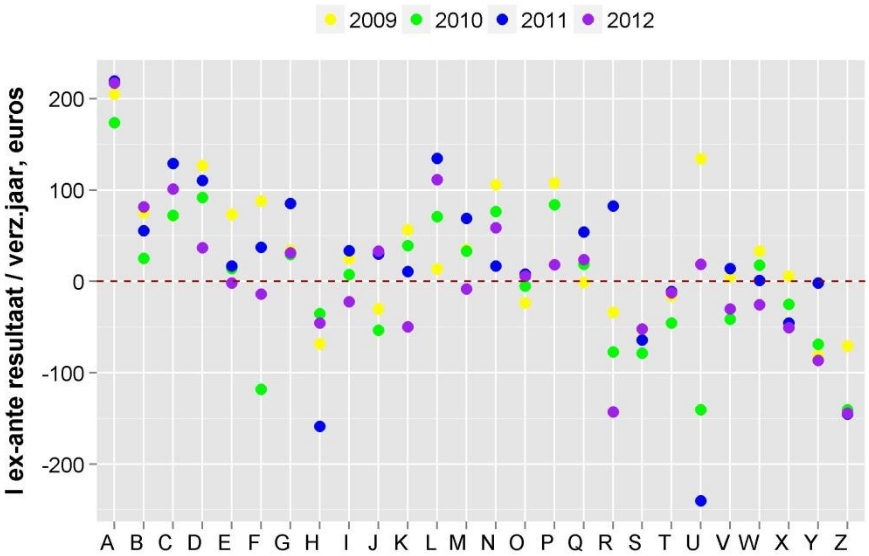
Figuur 5.3. Vereveningsresultaat op uitstroom



Figuur 5.4 laat het gemiddelde vereveningsresultaat per verzekerde zien voor de instroom per zorgverzekeraar. De spreiding van de resultaten tussen de zorgverzekeraars is duidelijk groter dan bij de vorige twee indicatoren. Dit wijst erop dat de samenstelling van de groep instromers per zorgverzekeraar meer divers is qua risicoprofiel dan de samenstelling van de totale populatie of de groep van uitstromers per zorgverzekeraar.[31] We zien dus eerder een selectieve instroom dan een selectieve uitstroom van verzekerden bij de zorgverzekeraars. Risicoselectie komt dus eerder tot stand doordat zorgverzekeraars verzekerden met een gunstig risicoprofiel aantrekken dan doordat zorgverzekeraars verzekerden met een ongunstig risicoprofiel proberen af te houden.

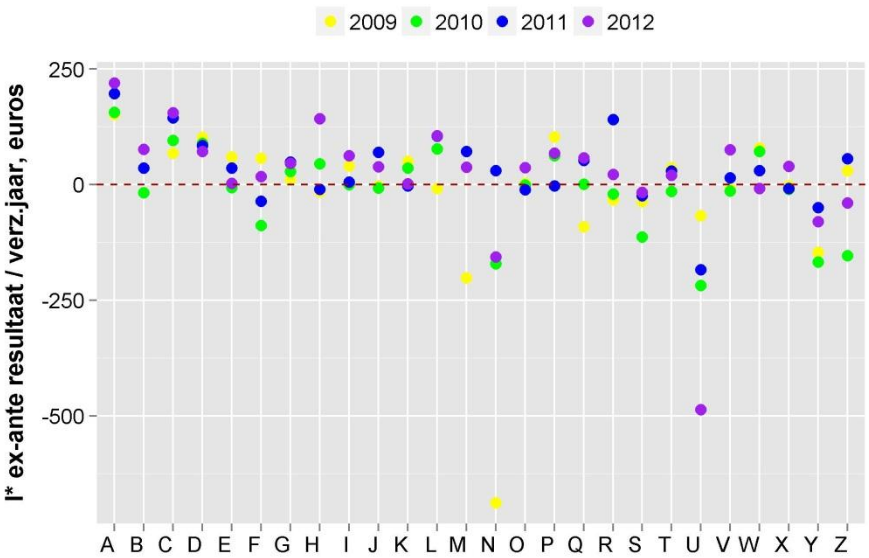
[31] De hele groep instromers en uitstromers is gelijk; beide zijn de groep van alle overstappers. Echter, als we naar de instroom en uitstroom per zorgverzekeraar kijken, zijn deze groepen niet gelijk (verzekerden die bij een zorgverzekeraar instromen zijn niet dezelfde verzekerden die in hetzelfde jaar deze zorgverzekeraar verlaten).

5.4. Vereveningsresultaat op instromers in jaar t



Zoals in Hoofdstuk 4 'Indicatoren' beschreven, berekenen we het vereveningsresultaat voor instromers van jaar t ook voor het jaar t-1 om (deels) voor doelmatigheid te corrigeren bij zorgverzekeraars van wie het resultaat op instromers positief is.
Figuur 5.5 laat zien dat het patroon van resultaten bij de meeste zorgverzekeraars niet verandert.

5.5. Vereveningsresultaat op instromers in jaar t-1



5.3 Effecten van niet voorspelbare kostenontwikkelingen per zorgverzekeraar

Om de indicator V te berekenen, wordt het econometrische vereveningsmodel opnieuw berekend op basis van het analyse bestand. Uit deze schatting komen de fictieve waarden van de 'normbedragen', die de daadwerkelijke allocatie van kosten over de kenmerken in kaart brengen.

Bij deze berekeningen hebben we ernaar gestreefd om dezelfde methodologie te gebruiken als bij de oorspronkelijke schatting van de normbedragen. Het deelbedrag 'Overig' wordt in de oorspronkelijke berekening van normbedragen met aparte modellen per kostentype (bijv. huisartsenzorg, verloskundige zorg) geschat. In onze dataset staat alleen het totaal van het deelbedrag 'Overig' zonder uitsplitsing naar type zorg. Hierdoor kunnen we de schattingen alleen op het totaal niveau doen, in plaats van het niveau van de oorspronkelijke deelmodellen.

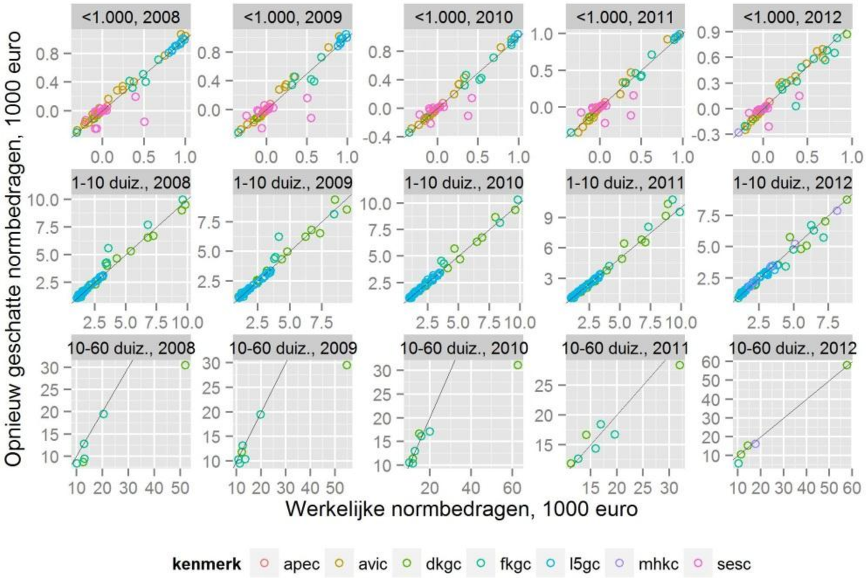
Voordat we de resultaten van indicator V per zorgverzekeraar laten zien, geven we een overzicht in hoeverre de werkelijke normbedragen overeenkomen met de opnieuw geschatte normbedragen. Risicoklassen waarbij de twee normbedragen ver uit elkaar liggen, worden geraakt door niet voorspelbare kostenontwikkelingen die niet meegenomen zijn in het ex-ante vereveningsmodel. Een andere verklaring voor het verschil tussen werkelijke en opnieuw geschatte normbedragen is dat een veronderstelling van het ex-ante model niet uitkwam in de werkelijkheid. In dit rapport geven we ter illustratie een overzicht van de resultaten. Een gedetailleerde analyse van de uitkomsten kan interessant zijn voor de bouwers van het ex-ante model.

In Figuur 5.6 zijn de werkelijke (horizontale as) en de opnieuw geschatte (verticale as) normbedragen tegen elkaar afgezet.[32] Elk normbedrag is een aparte cirkel waarbij de kleur van de cirkel naar het kenmerk verwijst waarbij het normbedrag hoort. De verschillende jaren zijn in aparte figuren in een rij weergegeven. In de eerste rij staan de figuren over de normbedragen tussen 0 en 1000 euro, in de tweede rij over de normbedragen tussen 1000 en 10.000 euro en in de derde rij die over de normbedragen boven 10.000 euro.

Uit de figuren blijkt dat de meeste werkelijke normbedragen goed overeenkomen met de opnieuw geschatte normbedragen. De uitschieters zijn risicoklassen waarin weinig verzekerden zitten waardoor de schattingen onstabiel worden.

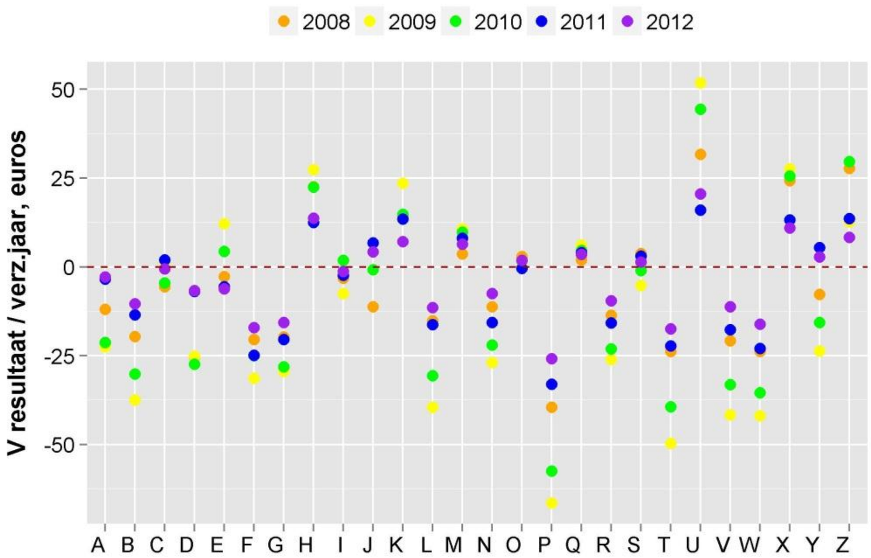
[32] In Figuur 5.6 zijn de nieuwe geschatte normbedragen gekalibreerd zodat de totale normatieve kosten over ons analysebestand overeenkomen met de totale normatieve kosten op basis van de werkelijke ex-ante normbedragen.

Figuur 5.6. Relatie werkelijke en opnieuw geschatte normbedragen



Figuur 5.7 laat het totale effect zien van de afwijking tussen werkelijke en opnieuw geschatte normbedragen per zorgverzekeraar. De meeste punten liggen tussen -25 en +25 euro. De spreiding is beperkter dan die van de indicator T, wat betekent dat de resultaten van indicator T wellicht beïnvloedt maar niet alleen bepaald zijn door niet voorspelbare kostenontwikkelingen.

Figuur 5.7. Indicator V per zorgverzekeraar



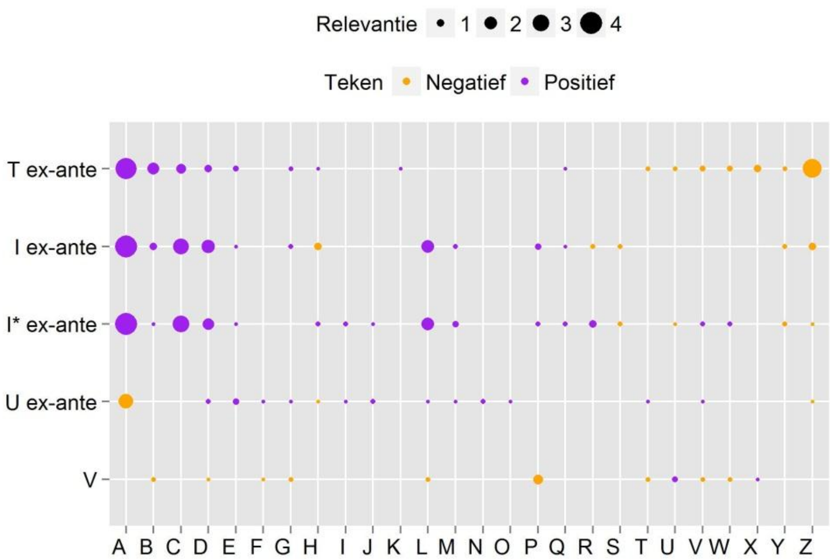
5.4 Overzicht per zorgverzekeraar

Om een overzicht te krijgen zetten we de resultaten van alle indicatoren per zorgverzekeraar in een grafiek. In Figuur 5.8. staan de verschillende indicatoren onder elkaar. Bij het beoordelen van de resultaten is het teken van de indicator (positief of negatief) belangrijk. Daarom onderscheiden we deze met kleuren.

Naast positieve of negatieve tekens, is ook het niveau van statistische significantie van belang. Bijvoorbeeld: bij een kleinere instroom is de kans dat het vereveningsresultaat toevallig erg hoog of erg laag uitvalt groter, want er kunnen toevallig een paar verzekerden tussen zitten van wie de waarde van het vereveningsresultaat extreem hoog of laag is. Bij een grotere instroom wordt dat minder aannemelijk. Dan levert de indicator ook meer betrouwbare informatie voor ons onderzoek, en is dan relevanter.

Om dit in kaart te brengen, kijken we naar de hoogte van de vereveningsresultaten. We onderscheiden vier positieve waarden: (1) *boven* 25 euro, (2) *boven* 50 euro, (3) *boven* 75 euro, en (4) *boven* 100 euro. En vier negatieve waarden: (-1) *onder* 25 euro, (-2) *onder* 50 euro, (-3) *onder* 75 euro, en (-4) *onder* 100 euro. Vervolgens toetsen we of de indicator deze benchmarkwaarden overschrijdt met een betrouwbaarheid van ten minste 95%. Bijvoorbeeld, als indicator T bij een zorgverzekeraar een positieve waarde heeft, toetsen we vier hypothesen om te bepalen of deze waarde statistisch significant boven de vier positieve niveaus uitkomt. Stel dat de hoogste hypothese, die we niet kunnen afwijzen, de hypothese (3) wordt. Dan zeggen we dat het relevantieniveau van deze indicator gelijk is aan 3. Dat geldt ook bij de negatieve waarden. Alleen wordt in dat geval de relevantie met negatieve waarden gemeten. De gemiddelde relevantie over een periode geeft ons dan een idee in hoeverre de indicatorwaarde afwijkt van nul.

Figuur 5.8. Samenvatting van alle resultaten per zorgverzekeraar



Figuur 5.8 geeft een samenvatting van de uitkomsten van alle indicatoren per zorgverzekeraar. In de eerste rij (indicator T ex-ante) links staan de zorgverzekeraars met een hoog relevante positieve vereveningsresultaat. Rechts staan zorgverzekeraars met een relevante

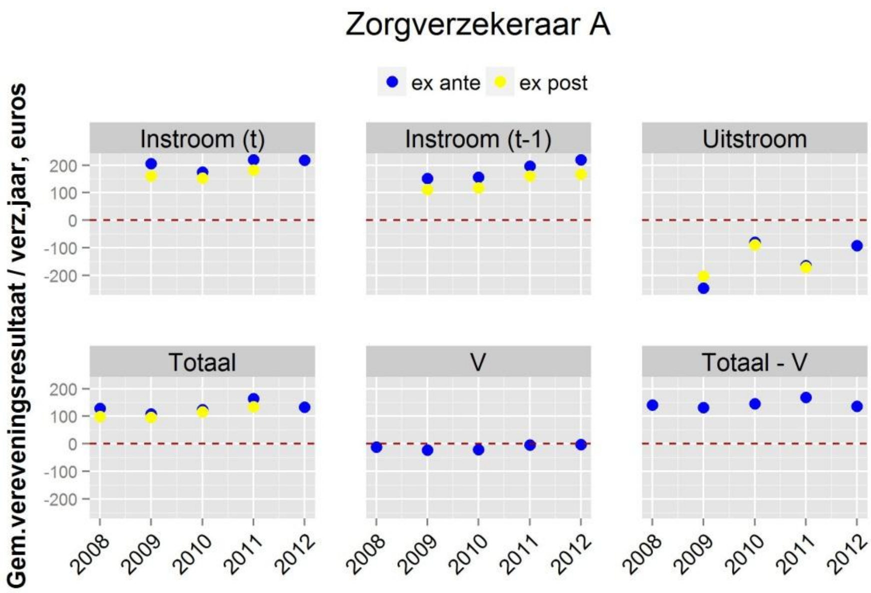
laag vereveningsresultaat. Het vereveningsresultaat op instroom (I ex-ante) toont een vergelijkbaar patroon. Bij de indicatoren U ex-ante en V zien we veel minder relevante resultaten.

Op basis van het overzicht van Figuur 5.8. kiezen we zorgverzekeraars A, C en Z als voorbeeld om de interpretatie van alle indicatoren in samenhang te laten zien. Zorgverzekeraar A en C hebben de meeste en hoogste significante resultaten. Zorgverzekeraar Z heeft een opvallend negatief resultaat op indicator T. Deze drie zorgverzekeraars analyseren we in de volgende paragraaf.

5.5 Case studies

In deze paragraaf analyseren we de uitkomsten van de geselecteerde gevallen uit Figuur 5.8. Het doel is om te laten zien hoe we de uitkomsten van verschillende indicatoren in samenhang interpreteren.

Figuur 5.9. Resultaten van Zorgverzekeraar A



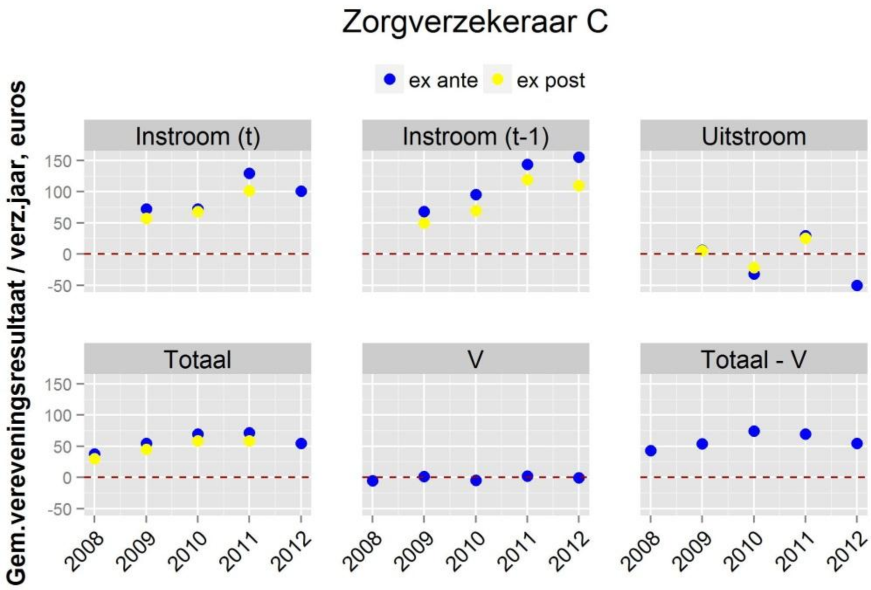
Figuur 5.9 vat de resultaten van Zorgverzekeraar A samen. Deze zorgverzekeraar heeft een structureel positief vereveningsresultaat per verzekerde op zijn totale populatie. Dit positieve resultaat kan echter nog het gevolg zijn van verschillende effecten. Het zou kunnen dat deze zorgverzekeraar vanuit oudsher een populatie met een gunstig risicoprofiel heeft. In dat geval zouden we ook een structureel positief gemiddeld resultaat op zijn uitstroom verwachten. Dat is echter structureel negatief. De verzekerden die Zorgverzekeraar A verlaten, hebben dus gemiddeld een ongunstig risicoprofiel. Het zou kunnen dat in de polisvoorwaarden van deze zorgverzekeraar iets staat waardoor het nadelig wordt voor verzekerden met een ongunstig risicoprofiel bij deze zorgverzekeraar te blijven. Deze hypothese testen we in het kader van het kwalitatief onderzoek.

De derde indicator, namelijk het gemiddelde resultaat over de instroom, bevestigt onze voorlopige bevindingen. Het gemiddelde resultaat op de instroom is structureel positief. Omdat het resultaat van instromers van jaar t ook in het jaar t-1 positief is, is het niet aannemelijk dat het gunstige resultaat door de doelmatigheid van de zorgverzekeraar komt. Verzekerden die bij Zorgverzekeraar A binnenstromen hebben structureel

een gunstig risicoprofiel. Ten slotte, indicator V laat zien dat onverwachte kostenontwikkelingen het vereveningsresultaat van Zorgverzekeraar A in beperkte mate beïnvloeden. Alle indicatoren samen geven een signaal voor risicoselectie, hoewel we die definitieve conclusie nog niet kunnen trekken in deze fase van het onderzoek.

Figuur 5.10 laat de resultaten van Zorgverzekeraar C zien.

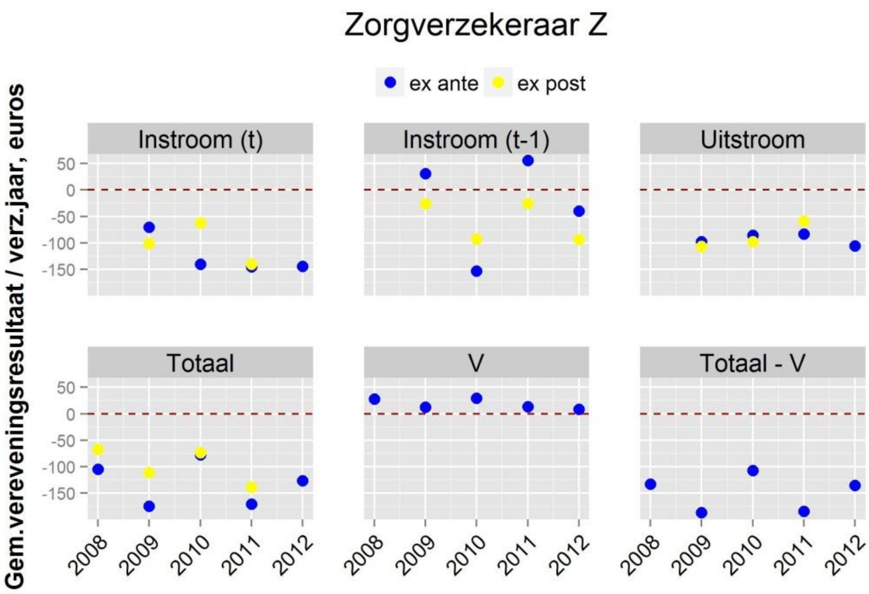
Figuur 5.10. Resultaten van Zorgverzekeraar C



Zorgverzekeraar C heeft ook een structureel positief vereveningsresultaat op zijn totale populatie. Het is wel lager dan dat van Zorgverzekeraar A. Het gemiddelde resultaat voor de uitstromers van Zorgverzekeraar C schommelt rond nul, wat de hypothese over een gunstige populatie vanuit oudsher niet bevestigt maar ook niet op risicoselectie wijst. Het gemiddelde resultaat op de instroom is structureel positief en wordt nog hoger als we voor doelmatigheid corrigeren. Dit betekent dat deze zorgverzekeraar verzekerden met een gunstig risicoprofiel weet aan te trekken. Verzekerden met een ongunstig risicoprofiel verlaten deze zorgverzekeraar echter niet in grote aantallen. Alles bij elkaar kan dit op risicoselectie wijzen, maar de aanwijzing is minder sterk dan bij Zorgverzekeraar A.

Figuur 5.11 hieronder laat de resultaten van Zorgverzekeraar Z zien.

Figuur 5.11. Resultaten van Zorgverzekeraar Z



Zorgverzekeraar Z heeft een structureel negatief gemiddeld resultaat op zijn totale populatie. Het resultaat op zijn uitstroom is ook structureel negatief. Dit kan er dus op wijzen dat deze zorgverzekeraar van oudsher een populatie met een ongunstig risicoprofiel heeft. Voor de instroom behaalt hij ook een negatief vereveningsresultaat. Dit laatste resultaat wordt in sommige jaren wel positief als we het voor het jaar t-1 uitrekenen. Dit kan erop wijzen dat de zorgverzekeraar ondoelmatig is, vergeleken met de gemiddelde zorgverzekeraar. Een andere verklaring is dat verzekerden juist op dat moment naar Zorgverzekeraar Z overstappen wanneer ze hogere zorgkosten verwachten. In jaar t is het vereveningsresultaat over deze verzekerden dus negatief terwijl in jaar t-1 het nog positief of nul was. De kostenontwikkelingen die niet in het ex-ante model waren meegenomen, hadden een licht positief effect op het vereveningsresultaat van deze zorgverzekeraar.

We kunnen echter nog niet definitief concluderen of er sprake is van risicoselectie bij deze zorgverzekeraars. De resultaten betreffen vooralsnog het vereveningsmodel voor de somatische zorg en nemen de effecten van het model voor de ggz en het eigen risico niet mee. Verder kan het niveau van de analyse, namelijk het niveau van de zorgverzekeraars (risicodragers), het beeld ook vertekenen. Resultaten van afzonderlijke polissen kunnen tegen elkaar uitmiddelen omdat we naar het hogere aggregatieniveau kijken.

6. Voorlopige conclusies en vervolg

De resultaten in het vorige hoofdstuk zijn de eerste inzichten in het bestaan van mogelijke risicoselectie op de zorgverzekeringsmarkt. Daaruit blijkt dat er structurele verschillen zijn tussen zorgverzekeraars als het gaat om het vereveningsresultaat per verzekerde op basis van het model voor de somatische zorg. De resultaten bevestigen het beeld dat in het rapport van iBMG 'Overstapgedrag en risicoselectie op de zorgverzekeringsmarkt' in 2013 werden geschetst op basis van de analyse van één overstapperiode (2008/ 2009).

Wanneer we de resultaten van alle indicatoren in samenhang analyseren, zien we signalen voor risicoselectie bij een paar zorgverzekeraars. We zien ook aanwijzingen dat risicoselectie eerder door het aantrekken van verzekerden met een gunstig risicoprofiel gebeurt dan door het afhouden van verzekerden met een ongunstig risicoprofiel. Deze voorlopige resultaten kunnen echter nog veranderen in de tweede fase van het onderzoek door het meenemen van de andere modellen van de risicoverevening (ggz en het eigen risico[33]) in onze analyses.

Zoals we in de inleiding schreven gebruiken we een brede definitie van risicoselectie, waar ook de zelfselectie van verzekerden onder valt. De signalen voor het bestaan van mogelijke risicoselectie die we op dit moment vinden, wijzen dus niet per se op een doelbewuste actie van de zorgverzekeraar om verzekerden met een gunstig risicoprofiel te selecteren. De signalen kunnen bijvoorbeeld ook het gevolg zijn van keuzes van de verzekerden. In die gevallen waar de signalen sterk zijn, zullen wij een kwalitatieve analyse uitvoeren om de gevonden resultaten te verklaren.

Tegelijkertijd zijn de gevonden signalen een onderschatting van de aanwezigheid van de breed gedefinieerde risicoselectie. Uit kwantitatieve gegevens zijn alleen de netto effecten van verschillende acties van zorgverzekeraars en verzekerden af te leiden. Dit betekent dat als zorgverzekeraars acties uitvoeren om risicoselectie onder verzekerden toe te passen die geen effect hebben, we die acties niet kunnen aantonen met kwantitatieve analyses. Ook voor deze reden zijn kwalitatieve methodes belangrijk.

Een andere belangrijke kanttekening bij de huidige analyses is dat we uitgaan van de werkelijk gemaakte kosten, die pas drie jaar na het onderzochte jaar beschikbaar zijn. Dit betekent dat de zorgverzekeraars niet over die cijfers beschikken op het moment dat zij hun strategie bepaalden. Verder, kennen zorgverzekeraars ook niet alle kenmerken van hun verzekerden. Ze hebben bijvoorbeeld geen informatie over de sociaal economische status en de aard van inkomen van hun verzekerden. Een zorgverzekeraar kan daarom het vereveningsresultaat per individuele verzekerde niet precies bepalen. Dat beperkt de mogelijkheid om deze informatie mee te nemen in hun (marketing)strategie.
Het resultaat dat we meten is dus niet alleen het gevolg van acties van zorgverzekeraars (en verzekerden), maar ook van externe factoren waarover niemand controle heeft. We nemen de gevolgen van eventuele onvoorspelbare (kosten)ontwikkelingen mee. De resultaten van de

[33] In de geanalyseerde periode (2008-2012) was het model Verpleging en Verzorging (V&V) nog niet van toepassing.

analyses met indicator V laten wel zien dat onvoorspelbare (kosten)ontwikkelingen maar een deel van het totale resultaat verklaren.

Bij het vergelijken van de resultaten is het ook van belang dat de analyses zijn uitgevoerd op het niveau van de zorgverzekeraars (risicodragers). Als een zorgverzekeraar meerdere polissen heeft, kunnen de resultaten van deze polissen elkaar uitmiddelen. Daarom zijn de resultaten van zorgverzekeraars met één polis niet vergelijkbaar met de resultaten van zorgverzekeraars met meerdere polissen. In het laatste geval onderschatten we mogelijk de signalen voor risicoselectie. Om deze reden doen we alle analyses in de tweede fase van het onderzoek op het niveau van polissen.

Zodra we de resultaten op het niveau van polissen kennen is het ook interessant om de uitkomsten van alle polissen binnen een zorgverzekeraarsconcern in samenhang te bekijken. Verzekerdenstromen binnen een concern kunnen ook licht werpen op strategieën van het concern. Het totale beeld dat we met concern-analyses krijgen, kan ons helpen om gedragingen van zorgverzekeraars(concerns) en mechanismen op de markt beter te begrijpen zodat bij ongewenste ontwikkelingen (indien mogelijk) doelgerichte maatregelen kunnen worden genomen.

In de tweede fase van het onderzoek willen we met statistische analyses het verschil in het vereveningsresultaat tussen polissen verklaren. Aan de hand van aanvullende gegevens zoeken we het statistische verband tussen het vereveningsresultaat van afzonderlijke polissen en de verschillende aanbodkenmerken van de polissen. Daarbij kijken we naar de polisvoorwaarden, de hoogte van de korting voor het vrijwillig eigen risico en naar collectieve contracten, zodat we de ontwikkelingen op de markt beter begrijpen en bij ongewenste ontwikkelingen maatregelen kunnen worden genomen.

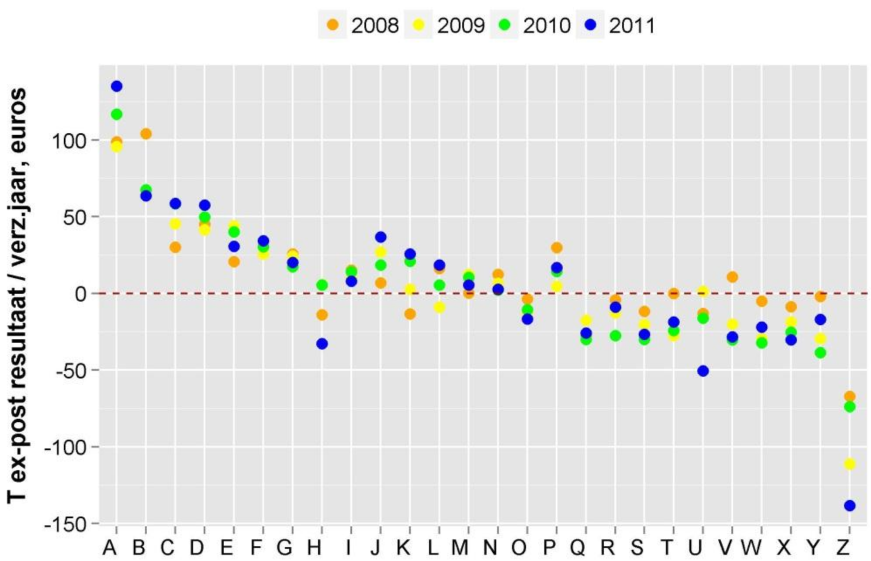
Het is echter niet mogelijk om op basis van kwantitatieve analyses ook de beweegredenen van zorgverzekeraars te begrijpen. Ook een beperking van onze analyse is dat de meest recente gegevens het jaar 2012 betreffen. Het is de vraag hoe de bevindingen uit de periode 2008-2012 relateren aan de huidige situatie. Om deze redenen begint de NZa met een kwalitatief onderzoek waarin verdiepend onderzoek wordt gedaan naar individuele casussen.

Bijlage 1. Aanvullende figuren

Hieronder staat de figuren van de ex-post resultaten van indicatoren T, U, I en I*. In de ex-post resultaten is het effect van band breedte regeling niet verwerkt. De reden hiervoor is dat de band breedte regeling op het niveau van zorgverzekeraars (risicodragers) van toepassing is. Indicatoren U, I en I* kunnen niet berekend worden inclusief de effect van band breedte regeling omdat deze een kleinere groep verzekerden betreffen dan die van de hele zorgverzekeraar.[34]

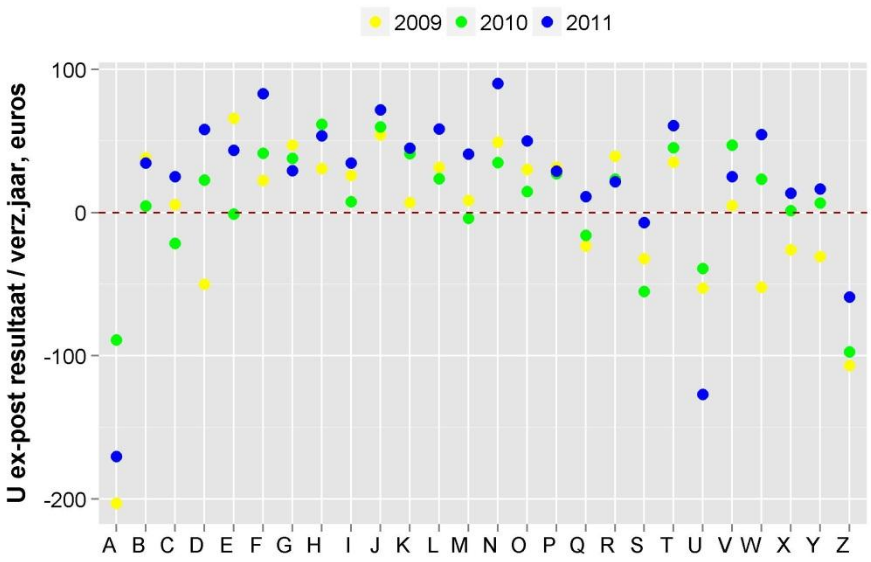
Het toepassen van bandbreedte-regeling op het vereveningsresultaat van de totale populatie van de zorgverzekeraar (indicator T) zou de extreme waardes van de indicator dichterbij nul brengen.

Gemiddelde ex-post resultaat op de totale populatie

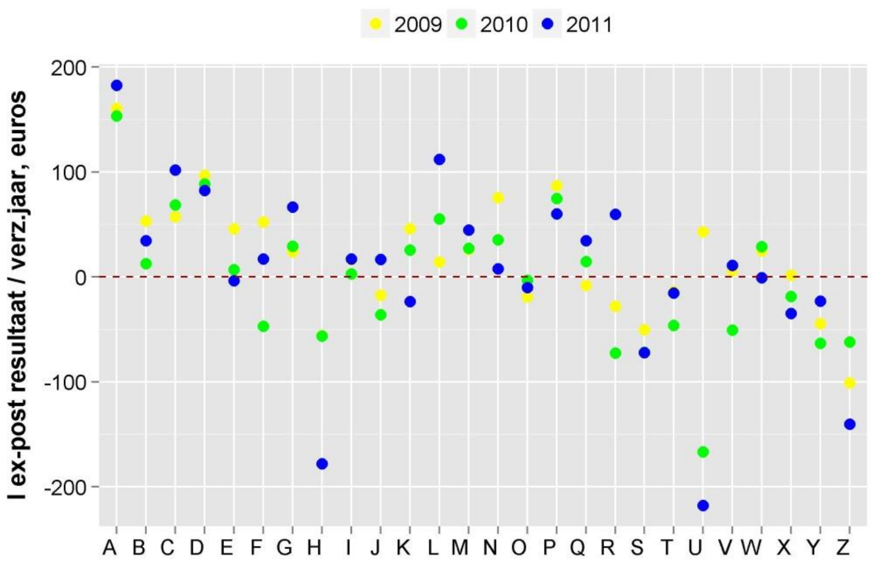


[34] Voor dezelfde redenen kan de band breedte regeling ook niet meegenomen worden in de tweede fase van het onderzoek waarin we de analyses op het niveau van polissen doen.

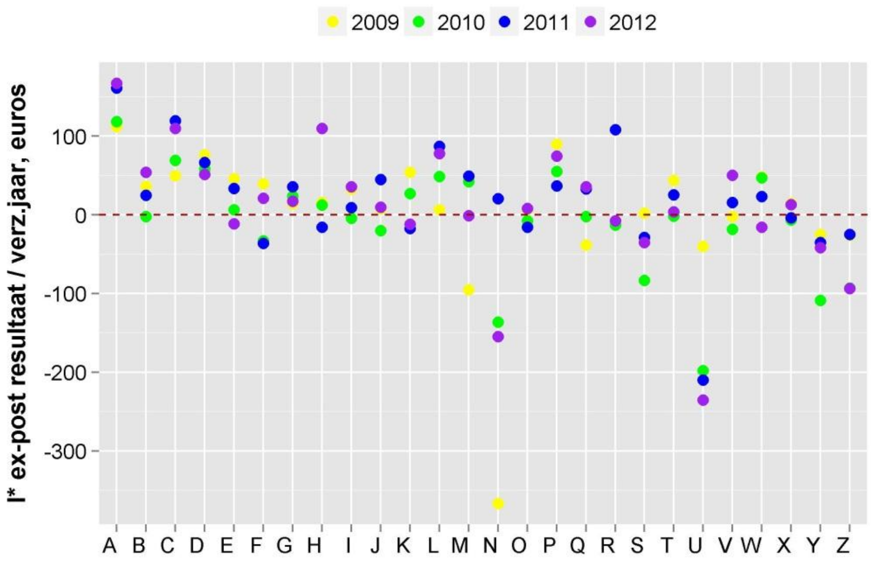
Gemiddelde ex-post resultaat op uitstroom



Gemiddelde ex-post resultaat op instroom jaar t



Gemiddelde ex-post resultaat op instroom jaar t-1



Bijlage 2. Technische toelichting

Deze bijlage beschrijft de berekeningsmethode en de aannames in de analyse. Dit betreft (1) de berekening van het ex-ante en ex-post vereveningsresultaat per verzekerde, (2) de keuzes die gemaakt zijn bij de samenstelling van het uiteindelijke bestand ten grondslag van de analyse, (3) de definitie van overstappers voor de analyse van de in- en uitstroom van verzekerden, (4) de benadering bij fusies, (5) uitleg over de regressie voor de berekening van de indicator V.

Hierbij nemen we het volgende als gegeven: het risicovereveningsmodel zelf, de werkelijke normbedragen en het originele verzekerdenbestand op het niveau van gepseudonimiseerd BSN-nummers dat Zorginstituut Nederland voor deze analyse beschikbaar heeft gesteld (hierna: 'het originele bestand').[35] Voor meer detail over de keuze van de variabelen voor het risicovereveningsmodel, de constructie van de dataset, en de berekening van de werkelijke normbedragen verwijzen we naar eerdere publicaties van Zorginstituut Nederland en instituut Beleid & Management Gezondheidszorg (Erasmus Universiteit Rotterdam) over deze onderwerpen.[36]

Berekening van het ex-ante en ex-post verevenings-resultaat op het niveau van verzekerden
Het vereveningsmodel van somatische zorg wordt op basis van drie kostenposten berekend: variabele kosten, DBC-kosten in het B-segment, en overige kosten.[37] Voor elke kostenpost geldt een aparte set van normbedragen. Deze worden gebruikt om de totale normatieve kosten per verzekerde te berekenen. Vervolgens wordt het resultaat per verzekerde berekend als het verschil tussen de normatieve en de werkelijke kosten. In ons onderzoek nemen we de normbedragen en werkelijke kosten als gegeven. Alle berekeningen worden per kostenpost gedaan. Het totale resultaat is de som van de resultaten over alle drie (of twee in 2008) type kosten.

De originele normbedragen die we als input gebruiken, zijn de definitieve normbedragen van Zorginstituut Nederland voor de berekening van kostenbudgetten (hierna 'ex-ante normbedragen') en de afrekenbudgetten (hierna 'ex-post normbedragen'). Deze originele normbedragen voldoen aan de voorwaarde dat het totale vereveningsresultaat per kostentype in elk jaar gelijk is aan nul. Om dit te bereiken past Zorginstituut Nederland de zogenoemde 'macronacalculatie' toe (een proportionele verhoging of verlaging van alle waarden om aan deze voorwaarde te voldoen).

[35] Dit bestand bevat alle verzekerden voor wie een BSN-nummer oorspronkelijk bekend was. Er is daarom een zeer kleine discrepantie in de aantallen verzekerdenjaren in ons bestand en in het bestand dat ten grondslag ligt van de definitieve vaststelling van Zorginstituut Nederland, waarin ook de verzekerden met een ontbrekende informatie over hun BSN-nummers zitten (bijv. in 2008 waren dit 500 verzekerdenjaren). Deze kleine discrepantie in aantallen heeft echter geen materiele gevolgen voor onze conclusies omdat het effect op de resultaten verwaarloosbaar is. Deze discrepantie wordt daarom in de rest van de bijlage genegeerd.
[36] O.a., Zorginstituut Nederland (2014) Handboek zorgverzekeraars informatie zorgverzekeringswet 2014; en iBMG (2014) Onderzoek risicoverevening 2015: berekening normbedragen, Onderzoek voor het Ministerie van Volksgezondheid, Welzijn en Sport, Eindrapportage, 24 september 2014, WOR 711.
[37] Daarnaast zijn er nog vaste kosten. De vaste kosten van alle verzekerden zijn volledig gecompenseerd. Daarom worden ze in onze analyse niet meegenomen.

De *normatieve ex-ante kosten* per kostenpost worden met *ex-ante normbedragen* berekend: (n^v, n^d, n^o). *Normatieve ex-post kosten* worden met *ex-post normbedragen* berekend (n^{*v}, n^{*d}, n^{*o}), waarbij elke letter n voor de respectievelijke vector van normbedragen staat, en de indices verwijzen respectievelijk naar variabele kosten (v), DBC kosten in het B-segment (d), en overige kosten (o). Door de normbedragen met de variabelen van risicoklassen te vermenigvuldigen, krijgen we de normatieve kosten per verzekerde. Hierbij telt elke verzekerde i met zijn gewicht w_i mee. Dit gewicht is een getal tussen 0 en 1 dat zegt hoelang de verzekerde bij een bepaalde zorgverzekeraar in dat jaar zat. Bijvoorbeeld, de normatieve ex-post variabele kosten van verzekerde i die een halfjaar bij een bepaalde zorgverzekeraar heeft gezeten, zijn gelijk aan $w_i n^{*v} X_i$, waarbij $w_i = 0.5$.

De risicoklasse-variabelen (notatie X_i hierboven) bestaan uit leeftijd- en geslachtsgroepen, farmaciekostengroepen, diagnosekostengroepen, regio's, sociale-economische status, de aard van inkomen en (vanaf 2012) meerjarige hoge-kosten groepen. De allocatie over risicoklassen binnen elke groep kan per jaar verschillen. Zie de beschrijving van gegevens in het rapport. Het totale aantal van alle kenmerkklassen is gelijk aan het totale aantal van normbedragen.

Het *ex-ante vereveningsresultaat* per kostenpost is het verschil tussen ex-ante normatieve kosten en werkelijke kosten van een verzekerde; en het *ex-post vereveningsresultaat* is het verschil tussen ex-post normatieve kosten en het aangepaste werkelijke kosten waarop enkele correcties worden toegepast. Bijvoorbeeld het ex-ante en ex-post vereveningsresultaten op variabele kosten van verzekerde i zijn als volgt berekend: $w_i n^v X_i - C_i^v$ en $w_i n^{*v} X_i - C_i^{*v}$. Gelijke formules gelden ook voor de andere twee kostenposten, waarbij de aangepaste kosten $(C_i^{*v}, C_i^{*d}, C_i^{*o})$ kleinere variatie hebben dan de werkelijke kosten (C_i^v, C_i^d, C_i^o).[38]

De berekening van aangepaste kosten bestaat uit enkele stappen. In onze analyse zijn alleen de correctiestappen van belang die het resultaat op het niveau van verzekerden kunnen beïnvloeden. Dit zijn: (1) macronacalculatie (jaren 2008-2011) of flankerend beleid (2012); (2) hoge kosten correctie (HKC); (3) nacalculatie van positieve en negatieve resultaten. De macronacalculatie – de correctie (1) dat vóór 2012 was toegepast – zorgt ervoor dat het totale vereveningsresultaat over alle verzekerden op nul uitkomt. Deze correctie is al van tevoren in de originele normbedragen verwerkt. De correcties (2) en (3) beïnvloeden verschillen tussen individuele verzekerden, en ze zijn daarom voor onze analyse van belang. Beide correcties verkleinen namelijk de relatieve verschillen tussen ex-post vereveningsresultaten van verzekerden in vergelijking met hun ex-ante vereveningsresultaten. Sinds 2012, is echter een extra kenmerk 'meerjarige hoge kosten' in het risicovereveningsmodel opgenomen. Sinds dat jaar worden eventuele verschillen betreffende hoge kosten via normbedragen gecompenseerd, en correcties (2) en (3) zijn daarom afgeschaft.[39]

Twee additionele correcties – namelijk de correctie voor de bandbreedteregeling (4) en de zorgkosten buitenland (5) – laten we in onze analyse buiten beschouwing. De bandbreedteregeling heeft in het

---

[38] Dit geldt voor alle jaren behalve het jaar 2008, waarin de kosten van DBC in het B-segment nog niet van toepassing zijn geweest.
[39] Op de vaste ziekenhuiskosten is er nog steeds 100% nacalculatie. Maar die nemen we niet mee in de analyses (en ze beïnvloeden de resultaten niet).

kader van deze studie geen betekenis want deze regeling heeft alleen op het niveau van een zorgverzekeraar (risicodrager) effect. De nacalculatie van curatieve zorgkosten buitenland via Zorginstituut Nederland– correctie (5) hierboven – betreft een relatief kleine groep van verzekerden. De achterliggende gegevens voor deze correctie zitten niet in de kostenbestanden op verzekerdenniveau; deze zijn daardoor niet in de analyse meegenomen. Het is aannemelijk dat deze correctie op het resultaat van verzekerden woonachtig in het buitenland een groter effect heeft. Door deze verzekerden uit het bestand te halen, corrigeren we het ontbreken van deze correctie in onze analyse.[40] Hierdoor beperken we de interpretatie van onze resultaten tot binnenlandse verzekerden.

De bovengenoemde HKC en nacalculatie worden als volgt berekend. De HKC wordt in gevallen toegepast waarin de som van de kosten van verzekerde i boven de HKC-benchmarkwaarde uitkomt. De totale HKC-inbreng van verzekerde i is gelijk aan het verschil tussen de kosten en de benchmarkwaarde, vermenigvuldigd met het percentage van de HKC:

$$I_i = \max(0, C_i^v + C_i^d + C_i^o - Benchmark_{HKC})\ Percetage_{HKC}$$

Dit bedrag wordt over de drie types kosten verdeeld. De totale inbreng per kostentype wordt daarna over alle verzekerden gelijk verdeeld. Verzekerden met hoge kosten worden zo gedeeltelijk gecompenseerd; dit wordt gefinancierd door van alle verzekerden vereveningsbijdrage in mindering te brengen.

Als voorbeeld laten we hier formules zien hoe de HKC correctie en nacalculatie voor variabele kosten wordt berekend. Dezelfde formules gelden ook voor de andere kostentypes. Het variabele kostenonderdeel van de HKC-inbreng wordt als volgt berekend:

$$I_i^v = \frac{C_i^v}{C_i^v + C_i^d + C_i^o} I_i$$

Deze inbreng wordt over alle verzekerdenjaren uit het oorspronkelijke verzekerdenbestand gelijk verdeeld, en er wordt het landelijke percentage berekend dat elke zorgverzekeraar voor elke individuele verzekerde in zijn bestand moet bijdragen: $\sum_j I_j^v / \sum_j w_j n^{*v} X_j$, waarbij j over alle waarnemingen uit het originele bestand loopt.[41]

Het resultaat van de HKC-correctie wordt als volgt bepaald:

$$C_{HKC,i}^v = (C_i^v - I_i^v) + w_i n^{*v} X_i \frac{\sum_j I_j^v}{\sum_j w_j n^{*v} X_j}$$

---

[40] In 2008 was dit ongeveer 0.7% van alle verzekerdenjaren.

[41] Doordat de verzekerden die woonachtig zijn in het buitenland uit het bestand zijn gehaald, en door enkele kleine discrepanties tussen ons bestand en de definitieve vaststelling van Zorginstituut Nederland (namelijk, kleine discrepantie in de verdeling van verzekerdenjaren over klassen in 2008, en correctie voor de zorgkosten buitenland) is onze landelijke percentage niet precies hetzelfde als het percentage waarmee het Zorginstituut definitieve bedragen berekent. Het verschil is echter in decimalen van procentpunten.

Na de HKC-correctie wordt nog nacalculatie toegepast, waarmee verschillen tussen resultaten van individuele verzekerden verder worden verkleind. Dit werkt als volgt:

$$w_i n^{*v} X_i - C_i^{*v} = (\, w_i n^{*v} X_i - C_{HKC,i}^{v})(1 - Percentage_{NAC}^{v})$$

Uit deze formule volgt de formule van aangepaste kosten per verzekerde:

$$C_i^{*v} = Percentage_{NAC}^{v}(\, w_i n^{*v} X_i) + (1 - Percentage_{NAC}^{v}) C_{HKC,i}^{v}$$

Tabel 8.1 vat de gebruikte parameters van de HKC en nacalculatie samen. De afnemende percentages van nacalculatie van variabele kosten in de tijd (zie de vierde kolom) betekenen dat het verschil tussen ex-ante en ex-post vereveningsresultaten van verzekerden steeds kleiner wordt. De percentages van 0 in 2012 betekenen dat de aangepaste kosten in 2012 gelijk zijn aan de werkelijke kosten, en dus dat het ex-ante en ex-post resultaat gelijk zijn.

Tabel 8.1. Parameters van HKC en nacalculatie

Jaar	Benchmark-Waarde HKC (EUR)	Percentage HKC	Percentage nacalculatie variabele kosten	Percentage nacalculatie B-dbc kosten	Percentage nacalculatie overige kosten
2008	20000	90	50	n.v.t.	0
2009	20000	90	40	15	0
2010	22500	90	30	0	0
2011	22500	90	30	0	0
2012	n.v.t.	0	0	0	0

Analysebestand van verzekerden

De bovenbeschreven methode wordt in onze studie toegepast om ex-ante en ex-post vereveningsresultaten per verzekerde te berekenen. Zoals hierboven uitgelegd bevat het uiteindelijke analysebestand de waarnemingen uit het originele gegevensbestand van Zorginstituut Nederland[42] over verzekerden woonachtig in Nederland. De waarnemingen over de woonachtigen in het buitenland[43] worden uit het dataset gehaald, omdat we de correctie voor deze groep verzekerden op basis van de beschikbare gegevens niet kunnen uitvoeren. Hierdoor beperken we de interpretatie van onze resultaten tot binnenlandse verzekerden. In 2008 was bij de meeste zorgverzekeraars het percentage van buitenlandse verzekerden kleiner dan 1.5%, alleen bij twee zorgverzekeraars was dit percentage hoger.

Het feit dat niet alle waarnemingen en correcties uit de praktijk in onze analyse worden meegenomen, zorgt voor een zekere discrepantie tussen de totale normatieve kosten en totale werkelijke kosten van alle verzekerden in ons bestand. Het totale vereveningsresultaat in onze

[42] Zie voetnoot 41 voor meer detail over een subtiele verschil tussen het 'originele bestand' gebruikt in onze studie en het bestand dat ten grondslag ligt van definitieve vaststellingen van het Zorginstituut.

[43] Deze waarnemingen worden aan de hand van de kenmerk 'regio' geselecteerd: regioklasse 'woonachtig in het buitenland'.

berekening komt hierdoor niet op nul uit.[44] Deze discrepantie is meestal kleiner dan 1%, behalve in het jaar 2012, waarin de macronacalculatie niet meer van toepassing was. Om deze discrepantie over verzekerden te verdelen maken we de aanname dat deze discrepantie proportioneel naar normatieve kosten is. Onder deze aannames kunnen we dan deze discrepantie over de verzekerden verdelen zodat het totale resultaat over het bestand gelijk aan nul wordt. Deze aanpassing heeft hetzelfde effect op het resultaat als macronacalculatie.

$$\sum_i (w_i n^{*v} X_i - C_i^{*v}) = \Delta^v$$

$$\sum_i \Delta_i^v = \Delta^v, \text{ waarbij } \Delta_i^v = \delta^v (w_i n^{*v} X_i)$$

Uit de bovenstaande formules volgen de uitdrukking voor δ^v:

$$\delta^v = \frac{\sum_i (w_i n^{*v} X_i - C_i^{*v})}{\sum_i w_i n^{*v} X_i}$$

De ex-ante en ex-post resultaten van verzekerde i worden daarna als volgt berekend:

$$(1 - \delta^v)\, w_i n^{*v} X_i - C_i^{*v}$$

$$(1 - Percentage_{NAC}^v)((1 - \delta^v)\, w_i n^{*v} X_i - C_i^{*v})$$

Een alternatieve benadering is om deze discrepantie gelijk over de verzekerden te verdelen, maar de resultaten van beide correcties zijn sterk met elkaar gecorreleerd. Hier kiezen we voor de eerste benadering omdat die als een extra slag van macronacalculatie werkt.
Bij ex-post berekeningen passen we deze correctie na de HKC-correctie toe.

Selectie van overstappers voor de analyse over de in- en uitstroom
De definitie van overstappers is eenvoudig voor verzekerden die per 1 januari overstappen en het hele jaar bij dezelfde zorgverzekeraar zitten. De geboorte of sterfte gevallen nemen we in de analyse van de in- en uitstroom niet mee.

Een belangrijk aandachtpunt betreft verzekerden die gedurende het jaar naar een andere zorgverzekeraar overstappen. Deze verzekerden hebben dan meer dan één zorgverzekeraar tijdens hetzelfde jaar. Deze situatie komt het meest voor bij jongeren die vanaf hun achttiende hun zorgverzekeraar zelf mogen kiezen. Daarnaast zijn er ook gevallen waarin een verzekerde om andere redenen gedurende het jaar naar een andere zorgverzekeraar kan overstappen.[45]

Het is aannemelijk dat de verzekerden die gedurende het jaar van hun zorgverzekeraar zijn veranderd, niet weer aan het eind van hetzelfde jaar overstappen. Onder deze aanname kunnen we dan de richting van de overstap, met een hoge waarschijnlijkheid, vaststellen om deze gevallen ook in de analyse mee te nemen. We beschouwen de overstappers die gedurende het jaar t overstappen als onderdeel van de in- en uitstroom van het jaar t. De overstappers die gedurende het jaar

[44] Een kleine discrepantie tussen ex-ante en ex-post normatieve kosten en werkelijke kosten bestaat ook in de praktijk doordat de normbedragen om praktische redenen op eurocenten worden afgerond.
[45] In de huidige situatie is het afgesproken dat verzekeraars deze niet meer toelaten. De wet zorgt wel voor enkele uitzonderingen zoals de situatie wanneer iemand van baan wisselt en daarom uit een collectiviteit moet. Het is duidelijker als we deze door de wet geregelde uitzonderingen noemen.

t-1 overstappen, beschouwen we als onderdeel van de in- en uitstroom van het jaar t-1. Zo voorkomen we een dubbeltelling van dezelfde overstap.

Naast deze gevallen is er ook een klein aantal gevallen waarin een verzekerde meerdere keren tijdens hetzelfde jaar overstapt. Deze gevallen laten we buiten beschouwing, want we weten niet welke zorgverzekeraar de eerste was en welke zorgverzekeraar de volgende was.

Om de in- en uitstroom van het jaar t te selecteren passen we de volgende procedure toe:

- *Stap 1*: de dataset van het jaar t wordt met de dataset van het jaar t-1 gekoppeld op de *combinatie van het pseudobsn-nummer van een verzekerde en het uzovi-nummer* van de zorgverzekeraar. De waarnemingen met een match beschouwen we als blijvers die in de overstapperiode tussen het jaar t-1 en het jaar t bij dezelfde zorgverzekeraar zaten.
- *Stap 2*: de rest van beide originele datasets van het jaar t en het jaar t-1 worden geschoond van duplicaten. Dit zijn een zeer klein deel van waarnemingen waarin de verzekerde (naast de zorgverzekeraar die in stap 1 al uitgehaald) tenminste bij twee andere zorgverzekeraars in hetzelfde jaar was verzekerd. Zoals hierboven is uitgelegd, nemen we deze waarnemingen in de analyse niet mee. Deze waarnemingen worden uit de datasets gehaald.
- *Stap 3*: de twee geschoonde datasets zijn aan elkaar gekoppeld *op het pseudobsn-nummer*. De waarnemingen met een match zijn verzekerden die van zorgverzekeraar zijn veranderd. Voor deze verzekerden wordt er daarna gecontroleerd of ze niet in het bestand van blijvers ook voorkomen. Als een verzekerde ook in het bestand van blijvers aanwezig is, dan heeft deze verzekerde twee zorgverzekeraars in beide jaren. En het is hoogstwaarschijnlijk dat deze verzekerde per 1 januari bij de zorgverzekeraar uit de dataset van blijvers zat. Deze verzekerden verschuiven we naar de volgende stap. De waarnemingen die overblijven zijn de overstappers die we in de analyse nemen.
- *Stap 4*: in deze stap vinden we de gevallen wanneer de verzekerde in het jaar t-1 en ook in een deel van het jaar t bij dezelfde zorgverzekeraar zat, maar midden het jaar t naar een andere zorgverzekeraar is overgestapt. Om deze gevallen te selecteren, wordt de rest van de geschoonde dataset t op de dataset van blijvers (van stap 1) gekoppeld. Degene met een match zouden ergens gedurende het jaar t-1 moeten zijn overgestapt. We beschouwen hun ook als instroom/uitstroom. We nemen dan in de analyse mee de waarneming van het jaar t, en de waarneming van het jaar t-1 wanneer de verzekerde nog bij andere verzekerde zat. De waarnemingen zonder match worden uit dataset gehaald.
- *Stap 5*: alle overstappers worden in dezelfde dataset samengevoegd.

Tabel 1 hieronder geeft de totale aantallen weer van verzekerdenjaren in elke stap van de selectieprocedure.

Tabel 8.2. Aantallen verzekerdenjaren bij selectie overstappers (in mln verzekerdenjaren)

	Stap 1			Stap 2		Stap 3	Stap 4
jaar	Gekoppeld	Niet gekoppeld jaar t-1	Niet gekoppeld jaar t	Uit-gehaald in jaar t-1	Uit-gehaald in jaar t	Overstap-pers uit deze stap	Overstap pers uit deze stap
2009	15,590	0,667	0,698	0,006	0,003	0,503	0,027
2010	15,590	0,742	0,781	0,004	0,004	0,596	0,026
2011	15,432	0,976	1,031	0,004	0,005	0,833	0,027
2012	15,392	1,098	1,132	0,005	0,005	0,946	0,027

Benadering bij fusies
In het geval van fusies beschouwen we de gezamenlijk in- en uitstroom van de gefuseerde bedrijven als de in- of uitstroom van het geïntegreerde bedrijf. De verzekerdenstromen tussen de twee bedrijven compenseren elkaar en ze beïnvloeden dan de berekening van indicatoren niet. Technisch gezien wordt dat gedaan door de aanpassing van de uzovi-code van het jaar (t-1), zodat de codes van beide bedrijven in beide jaren (t-1 en t) dezelfde worden. De interne overstappen vallen dan in de categorie 'blijvers' (stap 1 hierboven).

Modellen achteraf voorspellen
Gebruikmakend van het verzekerdenbestand met alle verzekerdenkenmerken die gekoppeld is met de werkelijke kosten per verzekerde schatten we de vereveningsmodel opnieuw.

Voor elk jaar en kostenpost schatten we het volgende lineaire model:

$$C_i = X_i\beta + \epsilon_i$$

X_i bevat de leeftijd- en geslachtsgroepen, farmaciekostengroepen (FKG), diagnosekostengroepen, regio's, sociale-economische status (SES), de aard van inkomen en (vanaf 2012) meerjarige hoge-kosten groepen. Voor elke categorie in elk groep geldt dat het als een aparte dummy is gecodeerd. Het model houdt rekening met gewichten van verzekerden w_i.

Bij het schatten van dit model passen we dezelfde restricties toe die ook bij het schatten van de werkelijke normbedragen werden toegepast. Dit zijn twee bindende restricties (op de FKG-risicogroep en voor een kleinste risicoklasse in de SES-risicogroep. Zie de definities in Hoofdstuk 3) en niet-bindende restricties op andere risicogroepen. Deze niet-bindende restrictie zorgen voor de zogenoemde 'criteriumneutraliteit' en voor de gelijke verdeling van de vereveningscompensaties over leeftijdscategorieën. Zie voor meer details in de studies in voetnoot 36.

Dit model schatten we per kostenpost apart: variabele kosten (v), DBC kosten in het B-segment (d), en overige kosten (o). In het originele vereveningsmodel is de kostenpost 'overige kosten' opgesplitst in verschillende categorieën, waarbij het lineaire model voor elke kostencategorie apart is geschat. Omdat we niet beschikken over de deelcategorieën van kostenpost overige kosten, schatten we de kostenpost 'overige kosten' alleen op totaal niveau.